TABLE 15: 169 GENES WITH SEQUENCE INFORMATION DEPICTED IN TABLE 16

Table 15 depicts UnigeneID, UnigeneTitle, Primekey, Predicted Cellular Localization, and Exemplar Accession for all of the sequences in Table 16. The information in Table 15 is linked by EosCode to Table 16.

Pkey:	Unique Eos probeset identifier number
ExAccn:	Exemplar Accession number, Genbank accession number
UnigeneID:	Unigene number
Unigene Title:	Unigene gene title
EosCode:	Internal Eos name
Localization:	Predicted cellular localization of gene product

Pkey	ExAccn	UnigeneID	Unigene Title	EosCode	Localization
100394	D84276	Hs.66052	CD38 antigen (p45)	PBC1	plasma membrane
100452	D87742	Hs.241552	KIAA0268 protein	PAB7	not determined
101249	L33881	Hs.1904	protein kinase C, iota	OAA1	cytoplasmic
101485	M24736		selectin E (endothelial adhesion molecul	ACC5	plasma membrane
101514	M28214	Hs.123072	RAB3B, member RAS oncogene family	PFJ2	cytoplasmic
101851	M94250	Hs.82045	midkine (neurite growth-promoting factor	LBH9	secreted
102398	U42359		gb:Human N33 protein form 1 (N33) gene,	PDG3	
102522	U53347	Hs.183556	solute carrier family 1 (neutral amino a	PFJ4	plasma membrane
102669	U71207	Hs.29279	eyes absent (Drosophila) homolog 2	LEM9	cytoplasmic
103119	X63629	Hs.2877	cadherin 3, type 1, P-cadherin (placenta	LBG2	plasma membrane
103709	AA037316	Hs.13804	hypothetical protein dJ462O23.2	PDO6	
104080	AA402971	Hs.57771	kallikrein 11	PBA6	secreted
104144	AA447439	Hs.183390	hypothetical protein FLJ13590	PDM3	
104691	AA011176	Hs.37744	Homo sapiens beta-1 adrenergic receptor	PAV1	plasma membrane
105370	AA236476	Hs.22791	transmembrane protein with EGF-like and	PDM9	plasma membrane
106149	AA424881	Hs.256301	hypothetical protein MGC13170	PDO8	
106579	AA456135	Hs.23023	ESTs	PAA4	plasma membrane
107102	AA609723	Hs.30652	KIAA1344 protein	PAA3	not determined
107217	D51095		DKFZP586E1621 protein	PDG8	
108153	AA054237	Hs.40808	ESTs	PBF1	plasma membrane
109014	AA156790	Hs.262036	ESTs, Weakly similar to Z223_HUMAN ZINC		PDG7
109112	AA169379	Hs.257924	hypothetical protein FLJ13782	BCU4	not determined
109890	H04649	Hs.20843	Homo sapiens cDNA FLJ11245 fis, clone PL		PDG4
110151	H18836	Hs.31608	hypothetical protein FLJ20041	PAV9	plasma membrane
112971	T17185	Hs.83883	transmembrane, prostate androgen induced		CHA1 not determined
113021	T23855	Hs.129836	KIAA1028 protein	PDO3	
114908	AA236545	Hs.54973	cadherin-like protein VR20	PFJ6	plasma membrane
114965	AA250737	Hs.72472	ESTs	BCY2	mitochondrial
116393	AA599463		hypothetical protein MGC2648	PDV3	secreted
116416	AA609219	Hs.39982	ESTs	OAB6	
117698	N41002	Hs.45107	ESTs	PDT9	ER
117984	N51919	Hs.106778	ATPase, Ca++ transporting, type 2C, memb		PAJ5 not determined
118985	N94303	Hs.55028	ESTs, Weakly similar to I54374 gene NF2	PDM8	
119018	N95796	Hs.278695	Homo sapiens prostein mRNA, complete cds		PAB2 plasma membrane
119126	R45175	Hs.117183	ESTs	PBF8	
120992	AA398246	Hs.97594	KIAA1210 protein	PDG5	
121710	AA419011		prostate androgen-regulated transcript 1	PDV5	
121913	AA428062		ESTs; protease inhibitor 15 (PI15)	BCU7	vesicular
122041	AA431407	Hs.98732	Homo sapiens Chromosome 16 BAC clone CIT		PAZ1 not determined
122593	AA453310	Hs.128749	alpha-methylacyl-CoA racemase	PDO1	
123209	AA489711	Hs.203270	ESTs, Weakly similar to ALU1_HUMAN ALU S		PAA2 plasma membrane
124526	N62096	Hs.293185	ESTs, Weakly similar to JC7328 amino aci	PAV4	plasma membrane
126399	AA128075		transmembrane, prostate androgen induced		PDY4
126645	AI167942	Hs.61635	six transmembrane epithelial antigen of	PAA5	plasma membrane
126966	R38438	Hs.182575	solute carrier family 15 (H+/peptide tra	PDO5	plasma membrane
127537	AA569531	Hs.162859	ESTs	PAA6	not determined
128790	AA291725	Hs.105700	secreted frizzled-related protein 4	BCX2	secreted
129109	AA491295	Hs.108708	calcium/calmodulin-dependent protein kin	PFJ7	
129184	W26769	Hs.109201	CGI-86 protein	PAV6	vesicular
129389	AA621604		spondin 2, extracellular matrix protein	CJA5	not determined

129404	AA172056		ESTs	PAB4	
129534	R73640	Hs.11260	hypothetical protein FLJ11264	PAJ3	secreted
130760	AA128997	Hs.18953	phosphodiesterase 9A	PEE6	nuclear
131425	AA219134	Hs.26691	ESTs	PBA7	
132964	AA031360		ESTs	PAA7	plasma membrane
132967	AA032221	Hs.61635	six transmembrane epithelial antigen of	PM17	plasma membrane
133179	U81599	Hs.66731	homeo box B13	PFJ5	nuclear
133330	U42360	Hs.71119	Putative prostate cancer tumor suppresso	PDM1	plasma membrane
133520	X74331	Hs.74519	primase, polypeptide 2A (58kD)	PDM2	
133724	U07919	Hs.75746	aldehyde dehydrogenase 1 family, member		PDT1 mitochondrial
133724	U07919	Hs.75746	aldehyde dehydrogenase 1 family, member		PDT1 mitochondrial
133944	AA045870	Hs.7780	Homo sapiens mRNA; cDNA DKFZp564A072 (fr		PAB9 cytoplasmic
134110	U41060	Hs.79136	LIV-1 protein, estrogen regulated	BCR4	plasma membrane
301805	AI800004	Hs.142846	hypothetical protein	PEU4	nuclear
302005	AI869666	Hs.123119	MAD (mothers against decapentaplegic, Dr	PBJ6	cytoplasmic
302881	AA508353	Hs.105314	relaxin 1 (H1)	PBH3	secreted
303506	AA340605	Hs.105887	ESTs, Weakly similar to Homolog of rat Z	PEG4	
303699	D30891	Hs.19525	hypothetical protein FLJ22794	PBM4	not determined
303753	AW503733	Hs.9414	KIAA1488 protein	PBY3	not determined
308050	AI460004	Hs.31608	hypothetical protein FLJ20041	PEU5	plasma membrane
310382	AI734009	Hs.127699	KIAA1603 protein	PCQ8	
310431	AI420227	Hs.149358	ESTs, Weakly similar to A46010 X-linked	PBH1	plasma membrane
310573	AW292180	Hs.156142	ESTs	PEN3	plasma membrane
310598	AI338013	Hs.140546	ESTs	PCW3	
310816	AI973051	Hs.224965	ESTs	PET5	
311596	AI682088	Hs.79375	holocarboxylase synthetase (biotin-[prop	PBH8	
313676	AA861697	Hs.120591	ESTs	PBY2	
314121	AI732100	Hs.187619	ESTs	PBY1	
314691	AW207206	Hs.136319	ESTs	BFF8	not determined
314785	AI538226	Hs.32976	guanine nucleotide binding protein 4	CBO7	cytoplasmic
314907	AI672225	Hs.222886	ESTs, Weakly similar to TRHY_HUMAN TRICH		PBM2 not determined
315051	AW292425		ESTs	PBM9	
315052	AA876910	Hs.134427	ESTs	PBJ7	plasma membrane
316442	AA760894	Hs.153023	ESTs	PBJ9	
317548	AI654187	Hs.195704	ESTs	PBQ6	
317869	AW295184	Hs.129142	deoxyribonuclease II beta	PBQ7	
318524	AW291511	Hs.159066	hypothetical protein FLJ10188	PBJ1	cytoplasmic
319191	AF071538		prostate epithelium-specific Ets transcr	PEN1	
319763	AA460775	Hs.6295	ESTs, Weakly similar to T17248 hypotheti	PEO7	
320324	AF071202	Hs.139336	ATP-binding cassette, sub-family C (CFTR	PBH5	plasma membrane
320561	NM_006953	Hs.159330	uroplakin 3	PEL9	plasma membrane
320796	AF038966	Hs.31218	secretory carrier membrane protein 1	PBY4	not determined
321441	AW297633	Hs.118498	Homo sapiens LUCA-15 protein mRNA, splic		PBY8 not determined
322303	W07459	Hs.157601	ESTs	CBF9	secreted
322782	AA056060	Hs.202577	Homo sapiens cDNA FLJ12166 fis, clone MA		PBQ1 not determined
322818	AW043782	Hs.293616	ESTs	PCQ7	plasma membrane
323226	AF055019	Hs.21906	Homo sapiens clone 24670 mRNA sequence		PCI2 not determined
323287	AA639902	Hs.104215	ESTs, Moderately similar to SPCN_HUMAN S	PBJ5	
324295	AI146686	Hs.143691	ESTs	PBQ9	not determined
324430	AA464018	Hs.184598	Homo sapiens cDNA: FLJ23241 fis, clone C		PBY6 not determined
324603	AW016378	Hs.292934	ESTs	PBM3	
324617	AA508552	Hs.195839	ESTs, Weakly similar to I38022 hypotheti	PBH4	cytoplasmic
324626	AI685464		gb:tt88f04.x1 NCI_CGAP_Pr28 Homo sapiens		PCW6
324658	AI694767	Hs.129179	Homo sapiens cDNA FLJ13581 fis, clone PL		PBJ4 plasma membrane
324718	AI557019	Hs.116467	small nuclear protein PRAC	CBK1	nuclear
330211				PBJ2	not determined
330546	U31382	Hs.299867	guanine nucleotide binding protein 4	PEW1	cytoplasmic
330762	AA449677	Hs.15251	hypothetical protein	PBM1	not determined
330790	T48536	Hs.122764	TMPRSS2, transmembrane protease, serine		PEL3 plasma membrane
330892	AA149579	Hs.91202	ESTs	PBQ4	plasma membrane
331099	R36671	Hs.14846	Homo sapiens mRNA; cDNA DKFZp564D016 (fr		PCQ1 cytoplasmic
331490	N32912	Hs.291039	ESTs	PCI4	nuclear
331889	AA431407	Hs.98802	ESTs, Moderately similar to T14342 NSD1	PBH7	not determined
332247	N58172		gb:za21f09.s1 Soares fetal liver spleen	PBQ5	nuclear
332396	AA340504		gb:hw31a09.x1 NCI_CGAP_Kid11 Homo sapien		PBJ8 not determined
332697	T94885		transgelin 2	PBQ8	secreted
332798				PBH2	nuclear
334447				PBY9	not determined
338255				PBY7	not determined

297

ID	Accession	Cluster	Description	Code	Location
401424				PFG2	mitochondrial
407122	H20276	Hs.31742	ESTs	PEW7	
408430	S79876	Hs.44926	dipeptidylpeptidase IV (CD26, adenosine	PEZ3	plasma membrane
408826	AF216077	Hs.48376	Homo sapiens clone HB-2 mRNA sequence		PEY1
409262	AK000631	Hs.52256	hypothetical protein FLJ20624	PFG1	nuclear
409361	NM_005982	Hs.54416	sine oculis homeobox (Drosophila) homolo	PEW3	nuclear
411096	U80034	Hs.68583	mitochondrial intermediate peptidase	PEZ9	mitochondrial
413125	BE244589	Hs.75207	glyoxalase I	PFJ3	cytoplasmic
413623	AA825721	Hs.246973	ESTs	OBH6	
414422	AA147224	Hs.337232	Homeo box A13	PFC6	
415263	AA948033	Hs.130853	ESTs	PEZ5	
417153	X57010	Hs.81343	"collagen, type II, alpha 1 (primary ost	PFJ1	secreted
418601	AA279490	Hs.86368	calmegin	PFA1	ER
418848	AI820961	Hs.193465	ESTs	PEY4	
418882	NM_004996	Hs.89433	ATP-binding cassette, sub-family C (CFTR	OBH2	
419839	U24577	Hs.93304	"phospholipase A2, group VII (platelet-a	PFH9	secreted
421887	AW161450	Hs.109201	CGI-86 protein	PFH2	plasma membrane
422083	NM_001141	Hs.111256	"arachidonate 15-lipoxygenase, second ty	PFH5	cytoplasmic
424565	AW102723	Hs.75295	guanylate cyclase 1, soluble, alpha 3	PFA3	
425071	NM_013989	Hs.154424	"deiodinase, iodothyronine, type II"	PFH6	secreted
425710	AF030880		solute carrier family, member 4	PFD4	plasma membrane
427958	AA418000	Hs.98280	potassium intermediate/small conductance	PFH1	plasma membrane
428819	AL135623	Hs.193914	KIAA0575 gene product	PFD6	nuclear
429900	AA460421	Hs.30875	ESTs	PEZ7	
429918	AW873986	Hs.119383	ESTs	PEY5	
430226	BE245562	Hs.2551	adrenergic, beta-2-, receptor, surface	PEZ4	plasma membrane
431217	NM_013427	Hs.250830	Rho GTPase activating protein 6	PFG6	nuclear
431716	D89053	Hs.268012	fatty-acid-Coenzyme A ligase, long-chain	PEZ1	
431992	NM_002742	Hs.2891	protein kinase C, mu	PFH4	cytoplasmic
432189	AA527941		gb:nh30c04.s1 NCI_CGAP_Pr3 Homo sapiens	PFA2	
432244	AI669973	Hs.200574	ESTs	PEW8	
432437	W07088	Hs.293685	ESTs	PFG3	
432966	AA650114	Hs.325198	ESTs	PEY3	
439176	AI446444	Hs.190394	ESTs, Weakly similar to B28096 line-1 pr	PEW5	
440260	AI972867	Hs.7130	copine IV	PEW6	
440901	AA909358	Hs.128612	ESTs	PFC8	
445424	AB028945		cortactin SH3 domain-binding protein	PEZ6	
446320	AF126245	Hs.14791	"acyl-Coenzyme A dehydrogenase family, m		PFH7
447210	AF035269		phosphatidylserine-specific phospholipas	PFH8	
449156	AF103907	Hs.171353	prostate cancer antigen 3, non-coding DD	PEZ8	
449625	NM_014253		odz (odd Oz/ten-m, Drosophila) homolog 1	PEZ2	plasma membrane
449650	AF055575	Hs.23838	calcium channel, voltage-dependent, L ty	PFD2	plasma membrane
451939	U80456	Hs.27311	single-minded (Drosophila) homolog 2	PFJ8	
451982	F13036	Hs.27373	Homo sapiens mRNA; cDNA DKFZp564O1763 (f		PFG9 plasma membrane
452039	AI922988		ESTs	PFD8	
452340	NM_002202	Hs.505	ISL1 transcription factor, LIM/homeodoma	PFG4	nuclear
452784	BE463857	Hs.151258	hypothetical protein FLJ21062	PFC5	cytoplasmic
452946	X95425	Hs.31092	EphA5	PFH3	plasma membrane

TABLE 15A shows the accession numbers for those primekeys lacking a unigeneID in Table 15. For each probeset we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Pkey: Unique Eos probeset identifier number
CAT number: Gene cluster number
Accession: Genbank accession numbers

Pkey	CAT number	Accession
116393	131543_1	AI972402 AI634409 AI523716 AI799749 W44518 AI424438 AI688513 AI971048 AI686324 AW013854 AA588483 AA528111 AI627428 AI582200 AI669296 AI826926 AI620526 AI669958 AI972458 AI924500 AA512903 W44517 AA335363 AW238997 BE300165 BE250665 AA284195 AA523420 W52834 AI471970 AI952824 AW003820 AW009463 AA669796 AA114966 AI653342 AA115038 AI342150 AI092100 AI968211 W51994 AI804005 AI201420 AI123210 AI738405 AI674964 AI970341 AW027500 AI493316 AI333193 AI139353 AA599463 AI656163 AI804200 AI365321 AI990213 AI657011 AA650025 AI968810 AI341978 AA599839 AW592602 AA644289 AI468578 AI565265 AI565228 BE221535 AW973052
101485	18113_1	AA296520 AL021940 M30640 NM_000450 M24736 M61894 AL047443 H39560 AI694691 AA916787 AI214796 AA939085 AI150616 AA412553 AA412545 AI051015 T27654 AA694430
126399	17331_1	AA088767 AF224278 AA128075 AL035541 AA027926 AI761441 AI972096 AW071693 AI742327 AI377498 AI804815 AI640802 AI885001 AI921394 AA595115 N71820 AI921217 AW007283 AI467828 AI369306 AA917446 AI493698 AA088701 AA126899 AI936228 AW204238 AI039567 AI925027 BE138909 AW452945 AW135998 AA310984 AA027860 AW073519 AI537597 AA953976 AI521341 AW273569 AW050740 AA536113 AA559064 AI474392 AW135709 AA535181 AW572959 AA570597 AI905464 AI677810 AI587642 AW975102 AA424310 AA482527 N64192 AA658276 AW889117 AA486591 AW889172 AI381990 AI381991 AI673419 AI990950 AA487031 AI272934 AI150565 AA229168 AW316722 AI142707 BE222396 AA614168 AA122026 AW338227 AA632457 AI968726 AW369662 AA512956 AA541675 AA451748 AI250993 BE146418 AA122025
132964	94346_1	AI362575 AI805082 AW263421 AI432462 AA135870 AA031360 AA031604 AA298475 AA298464
129389	21074_1	NM_012445 AB027466 BE407510 BE047605 AA047125 AW084003 AA149494 AA149490 AA292528 AA570505 AA526186 AW006250 AW007762 AI341557 AI799666 AI972710 AI377966 AI962810 AI084783 AI458032 AI190971 AW148913 AA372354 AW970032 AW007426 AA650188 AI123203 AI122890 AI280975 W73595 W73495 AI863238 AA374109 AA603986 AW149089 AW957523 AI307748 AI921067 AI336463 F24537 AI380460 AI367500 AI189309 AI814701 AI766921 AW572106 AA037024 AW072576 AA578293 AI288103 AA235464 AW450642 AA574230 AW294024 AI589229 AI580733 AW512227 AA877009 AI660255 AW188597 AA558228 AI572782 AA658397 AI274628 AI866359 AA864573 AI264439 AA621604 AW515493 AW243333 Z39737 AI567038 AA573997 AA573559 AW236431 AI652870 AI684973 AA034505 AA047126
129404	156454_1	AI267700 AI720344 AA191424 AI023543 AI469633 AA172056 AW958465 AA172236 AW953397 AA355086
107217	9836_1	AL080235 AA031750 D81382 AI480231 AI095947 AI560953 BE010721 AI870290 AA374945 AA125792 D51527 D51556 AI685541 D51559 AW117286 AA195741 AI675138 AW593439 AI201885 T30590 AW952100 D51095 AA523864 W70043 AA987586 AI421515 AI205532 AA127069 AI337367 D51595 AI453785 AW075677 AW088359 C14287 C14284
121710	19266_1	AF163474 NM_016590 AF163475 AI761105 AI770098 AA410580 AA411616 AI590343 AI739050 AL050198 AI862645 AA419104 AA513809 AA333032 AI816915 AW139625 AA640889 AI311391 AI627693 AW135514 AA419011 AI269149 AI245259 AI970008 AI970017 AW139445 AA569503 AI761072 AI766179 AI759995 AI300776 AI870129 AW150770 AA226501 AA226220
121913	291015_1	AI249368 AI742316 AA428062 AA442089 AI864189 BE349478 AI803475 AI584049 BE552085 AI088609 AI264197 AI886144 AI129474 AI307145 BE181300 AW058403 AI696838 AW748598 AA442196 AI216428
102398	entrez_U42359	U42359
315051	347217_1	AW292425 BE467167 AI702953 BE550961 BE222309 AI299348 AI693336 AA541708
324626	336411_1	AI685464 AW971336 AA513587 AA525142
319191	16065_1	NM_012391 AF071538 AB031549 AI685592 AI745526 AA662204 AW130657 AA662164 AW971121 AI668916 AA513274 AI991223 AI979170 AW298436 AA639821 AI859010 AW513942 AI687669 AA662521 AA548598 AI345056 AI305374 BE043418 AI432856 AI334840 AI379796 AI492693 AI307915 BE042082 AI307834 AI307858 AI309488 BE042210 AI435670 AI371605 AI862491 AI284563 AI306872 AI255044 AI254601 AI251236 AI473073 AI473042 AI432760 AI435664 AI336826 AI289365 AI369096 AI862274 AI334871 AI349863 AI250405 AI377617 AI309895 AI313017 AI862291 AI311936 AI378718 AI305722 AI306769 AI308888 AI334565 AI862296 AI344230 AI435685 AI344087 AI378696 AI311209 AI435775 AI310611 AI311154 AI432289 AI431561 AI492681 AI432867 AI335288 AI492796 AI432769 AI310299 AI432273 AI379820 AI275319 AI435753 AI609441 AI432767 AI369100 AI311420 AI349974 AI247157 AI334677 AI270910 AI224320 AI305608 AI334489 AI377152 AI350012 AI370086 AI335053 AI306781 AI306750 AI334849 AI334874 AI340380 AI307876 AI305974 AI305972 AI311521 AI334872 AI862509 AI311498 AI335051 AI289684 AI310859 AI311862 AI862483 AI492775 AI307906 AI492708 AI289693 AI340373 AI307910 AI311359 AI435653 AI334865 AI311492 AI492809 AI492690 AI431576 AI862268 AI311879 AI308435 AI492792 AI862512 AI275321 AI431568 AI431564 AI307885 AI307926 AI435692 AI435778 AI310182 AI308894 AI492707 AI492713 AI308560 AI307829 AI343234 AI580598 AW472796 AI340918 AI310243 AI309368 AI307920 AI289665

299

AI306777 AW086318 AW086292 AW086378 AI310027 AI275293 AI369082 AI340900 AI306749 AI371558 AW086287 BE043803
AI306793 AI306272 AI287948 AI270917 AI284816 AI336813 AI284546 AI308044 AI275290 AI270872 AI306795 AI289687 AI223570
AI305303 AI289677 AI287742 AI275284 AI306812 AI336701 AI371554 AI378719 AI344988 AI223631 AI335141 AI343222 AI284568
AI305357 AI275270 AI345932 AI436549 AI307925 AI311502 AI344238 AI343182 AI308508 AI305988 AI270790 AI379792 AI305647
AI305410 AI432251 AI436517 AI343227 AI305534 AI340387 AI271043 AI305499 AI271046 AI305962 AI289465 AI305378 AI289725
AI310848 AI305848 AI289362 AI252964 AI307049 AI310831 AI306993 AI306796 AI224659 AI305969 AI349855 AI306164 AI306948
AI284676 AI309155 AI343202 AI432785 AI306815 AI369081 AI270885 AI289699 AI435704 AI309647 AI305716 AI311281 AI287927
AI472995 AI340423 AI270958 AI307069 AI305364 AI270807 AI275306 AI311890 AI275263 AI432750 AI289371 AI432861 AI255113
AI305709 AI473008 AI311168 AI309711 AI377164 AI271201 AI289560 AI309710 AI306195 AI311201 AI287741 AI271066 AI432876
AI275281 AI379795 AI472972 AI311967 AI306826 AI305465 AI270792 AI473019 AI305340 AI270922 AI305995 AI305462 AI254144
AI270969 AI473012 AI305390 AI275278 AI223644 AI289692 AI250318 AI305372 AI289691 AI250521 AI306283 AI306814 AI307933
AI473160 AI432903 AI223720 AI254979 AI334862 AI306926 AI289541 AI432248 AI435722 AI435698 AI432859 AI310683 AI473175
AI335144 AI289467 AI436489 AI306928 AI473033 AI305763 AI307868 AI307882 AI348959 AI435736 AI432857 AI432896 AI435735
AI432283 AI473086 AI432863 AI473081 AI432825 AI307840 AI473164 AI432885 AI473166 AI472982 AI435734 AI473060 AI473171
AI432279 AI432882 AI334670 AI436512 AI432827 AI432852 AI473051 AI473077 AI435697 AI271509 AI492781 AI472983 AI473018
AI432897 AI473043 AI432871 AI436536 AI473157 AI349715 AI432777 AI473016 AI473158 AI340369 AI307941 AI432773 AI377146
AI492791 AI270950 AI305342 AI284604 AI306269 AI284811 AI270811 AI289347 AI334869 AI334852 AI311759 AI250382 AI309520
AI289550 AI305721 AI340870 AI270901 AI308575 AI307904 AI340715 AI270941 AI309808 AI246867 AI473014 AI307039 AI289360
AI473069 AI492786 AI344013 AI305876 AI436510 AI340742 AI473028 AI307891 BE041871 BE041268 BE042340 BE041946
BE041783 AI306173 AI201948 AI926972 AI275769

338255 CH22_6856FG__LINK_EM:AC00
330211 c_5_p2
332798 CH22_14FG_6_5_LINK_C4G1.G
334447 CH22_1746FG_387_7_LINK_EM
332247 372969_1   AA669097 AA513815 AA026798 AA676526 AA704429 AA704269 AW118292 AA579216 N58172
332396 20265_1   AW579842 BE156562 BE156690 BE156489 BE081033 AK001559 BE149402 M85387 AW367811 AW367798 R17370 AI908947
AA382932 R58449 H18732 AA371231 AW962899 AA713530 AW892946 R53463 H11063 AW068542 Z40761 BE176212 BE176155
W23952 W92188 AW374883 AA303497 AW954769 AA036808 BE168063 AW382073 AW382085 AL041475 H80748 AI078161
BE463983 AI805213 AI761264 W94885 N94502 AI623772 AI419532 AI810302 AI634190 AW002516 AW150777 AI352312 AI367474
AW204807 AI675502 AI337026 AW134715 BE328451 AI123157 AI560020 AI300745 AI608631 AI248873 AA742484 AW051635
H18646 AI245045 AA507111 AI640510 AI925594 AA115747 AA143035 AA151106
332697 13699_1   X51405 NM_001873 T11322 AL118886 BE328175 AW136009 BE467445 AW470313 AA774852 BE504139 AW501046 AA082792
AW389231 AA370044 R36841 AA371457 C04813 R25791 R25556 AW895854 AW903819 AW895671 AW895677 BE159723
AW895664 AW895597 AW895595 AW895665 AW888518 AI903724 F06081 F08503 AL119462 AW895730 AW888516 R26511
R26489 AA334126 AA327626 N85713 AW895998 AA223622 F05468 AA370749 W05590 M78202 AA371073 AW498607 R15017
T16991 AA001282 AA001138 AA551566 AA330159 AI922855 AA383512 AA029603 D82246 D82171 T94933 H56545 AA348060
AA176888 R96764 AW451817 AA385766 AA452618 AI690057 AA988822 BE549928 AA150901 W57992 AW899925 C05281
AA932042 AA370980 AW962877 W04741 AA369982 AW385948 AA922466 N75882 AI422070 AI361256 AI680224 D57122 T94885
R53266 R46713 T19071 AW796277 AA325333 F04719 F02334 AA358146 AA626597 AA358304 AW028099 AL119570 D57290
D58273 D57796 N48555 AI361969 AA329457 D57225 AW024046 AA992606 AW022118 AW021538 AA935845 H89870 H56546
AW961219 AA453239 AW837541 N45521 BE218029 AA318877 AA327740 AW961809 T92139 D53216 D52365 D53363 D53312
D53116 AI547267 AA679935 AW026552 AW190507 AI927710 AW244108 D50948 AW054991 AW021063 AW022511
AA493436 AI365636 BE464751 AW149384 AA102442 AW771368 AI818251 AI126368 D51049 AI421542 AI559467 AW079779
AW021048 AW023969 AW044214 AI458264 AA027274 AI620254 AW028917 BE219511 AA326242 N67561 AI971273 AA878328
D57131 AA770662 AI309299 AI796767 AA613338 W58076 AI566287 AI445573 AI880260 AA001919 AW339259 AI492610 AI492611
R97692 AI301425 AA722603 D58361 AI350323 AA973926 AI431263 AA516126 AA865467 AI925177 N39443 AA001943 AI299371
AI082412 AA665090 AA583433 H89871 AA977231 AI362219 AI056096 AI270446 N67524 N22103 AW614224 AA744054 AW243622
AI613188 AI929173 AI350243 AI362138 AA744004 AA176661 D56787 AI955625 AI393109 AI094769 AI479728 AI423107 AI955617
AI034036 AI582196 AW264534 AI418961 AA570761 AI343538 AA650341 AA992503 AA770004 AL039666 AI862675 AW190335
AA610274 AW418627 BE467472 D56786 T28749 AI217610 AI359556 T23523 AL040189 AA846222 AA651636 D51280 AI888986
AI521167 AI340177 AW612815 AI625285 AA621607 AA177059 AA229768 AA829788 AI749682 AW190631 N75299 AA230089
AI915632 BE069542 AA890020 AA528397 AA995390 BE503860 AA570812 AW339396 AI197986 AI203725 AI282379 AA670375
AA461513 F01728 AW243599 C00856 N75567 R95995 AA150932 R95961 AA648060 AA933800 AA927073 AA101126 AA864190
T93566 BE167472
425710 25529_1   AF030880 NM_000441 AC002467 AA385554 H23053 AW891838 AI139968 AA653057 AI695233
432189 342819_1   AA527941 AI810608 AI620190 AA635266
445424 6391_1   AB028945 T77648 F13328 AL157605 Z46212 AA304736 F11855 T66098 T30174 AW954164 AW176301 AW748243 AA456428
AI369958 AA938565 AW959613 Z42008 AA994779 AI683909 F11019 F10926 AI769597 AI752550 T65015 AI884314 AA643954
Z41838 AW020147 AI038822 AW571822 AA299781 AA894928 AF131790 BE005411 AI902476 AW082695 AA464384 R42750
AW902301 AA464273 R05837 Z38294 H41098 AL134507 M86079
447210 7119_1   AF035269 AF035268 NM_015900 T96213 U37591 AA156832 AA299371 AI084325 H95977 AI765967 BE221465 AA156726 AI969563
AW024539 AI436791 AI949451 AA843093 AI452756 AA824232 AI306667 T96131 AW207447 AW243556 AW957032 AI084332
H95978 U30998
449625 8113_1   NM_014253 AF100772 BE088769 AL022718 BE161779 AW863569 BE161640 AL039060 BE168542 AW296554 AA323193 AA235370
AW779760 N48674 AI375997 R45432 D59344 AI203107 F07491 R35360 R25094 AI913631 AI498402 T61382 AI016320 N45526
T61415 AA331486
452039 89513_1   AI922988 H05475 AA021608 AW169947 AA913750 Z41614 AW800012

TABLE 15B shows the genomic positioning for those primekeys lacking unigene ID's and accession numbers in Table 15. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also

5   listed.

Pkey:	Unique number corresponding to an Eos probeset
Ref:	Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA
	sequence of human chromosome 22." Dunham I. et al., Nature (1999) 402:489-495.
Strand:	Indicates DNA strand from which exons were predicted.
Nt_position:	Indicates nucleotide positions of predicted exons.

Pkey	Ref	Strand	Nt_position
334447	Dunham, I. et.al.	Plus	14308764-14308824
332798	Dunham, I. et.al.	Minus	232147-231974
338255	Dunham, I. et.al.	Minus	15242294-15242231
330211	6013592	Plus	59158-59215
401424	8176894	Plus	24223-24428

301

TABLE 11 AND SEQUENCE LISTING

SEQ ID NO:1 BCU4 DNA SEQUENCE

Nucleic Acid Accession #: NM_024915
Coding sequence: 13-1890 (underlined sequences correspond to start and stop codons)

```
1    11   21   31   41   51
|    |    |    |    |    |
ATTGGATCAA ACATGTCACA AGAGTCGGAC AATAATAAAA GACTAGTGGC CTTAGTGCCC  60
ATGCCCAGTG ACCCTCCATT CAATACCCGA AGAGCCTACA CCAGTGAGGA TGAAGCCTGG  120
AAGTCATACT TGGAGAATCC CCTGACAGCA GCCACCAAGG CCATGATGAT CATTAATGGT  180
GATGAGGACA GTGCTGCTGC CCTCGGCCTG CTCTATGACT ACTACAAGGT TCCTCGAGAC  240
AAGAGGCTGC TGTCTGTAAG CAAAGCAAGT GACAGCCAAG AAGACCAGGA GAAAAGAAAC  300
TGCCTTGGCA CCAGTGAAGC CCAGAGTAAT TTGAGTGGAG GAGAAAACCG AGTGCAAGTC  360
CTAAAGACTG TTCCAGTGAA CCTTTCCCTA AATCAAGATC ACCTGGAGAA TTCCAAGCGG  420
GAACAGTACA GCATCAGCTT CCCCGAGAGC TCTGCCATCA TCCCGGTGTC GGGAATCACG  480
GTGGTGAAAG CTGAAGATTT CACACCAGTT TTCATGGCCC CACCTGTGCA CTATCCCCGG  540
GGAGATGGGG AAGAGCAACG AGTGGTTATC TTTGAACAGA CTCAGTATGA CGTGCCCTCG  600
CTGGCCACCC ACAGCGCCTA TCTCAAAGAC GACCAGCGCA GCACTCCGGA CAGCACATAC  660
AGCGAGAGCT TCAAGGACGC AGCCACAGAG AAATTTCGGA GTGCTTCAGT TGGGGCTGAG  720
GAGTACATGT ATGATCAGAC ATCAAGTGGC ACATTTCAGT ACACCCTGGA GCCACCAAA  780
TCTCTCCGTC AGAAGCAGGG GGAGGGCCCC ATGACCTACC TCAACAAAGG ACAGTTCTAT  840
GCCATAACAC TCAGCGAGAC CGGAGGACAAC AAATGCTTCC GACACCCCAT CAGCAAAGTC  900
AGGAGTGTGG TGATGGTGGT CTTCAGTGAA GACAAAAACA GAGATGAACA GCTCAAATAC  960
TGGAAATACT GGCACTCTCG GCAGCATACG GCGAAGCAGA GGGTCCTTGA CATTGCCGAT  1020
TACAAGGAGA GCTTTAATAC GATTGGAAAC ATTGAAGAGA TTGCATATAA TGCTGTTTCC  1080
TTTACCTGGG ACGTGAATGA AGAGGCGAAG ATTTTCATCA CCGTGAATTG CTTGAGCACA  1140
GATTTCTCCT CCCAAAAAGG GGTGAAAGGA CTTCCTTTGA TGATTCAGAT TGACACATAC  1200
AGTTATAACA ATCGTAGCAA TAAACCCATT CATAGAGCTT ATTGCCAGAT CAAGGTCTTC  1260
TGTGACAAAG GAGCAGAAAG AAAAATCCGA GATGAAGAGC AGAAGCAGAA CAGGAAGAAC  1320
GGGAAAGGCC AGGCCTCCCA AACTCAATGC AACAGCTCCT CTGATGGGAA GTTGGCTGCC  1380
ATACCTTTAC AGAAGAAGAG TGACATCACC TACTTCAAAA CCATGCCTGA TCTCCACTCA  1440
CAGCCAGTTC TCTTCATACC TGATGTTCAC TTTGCAAGCT TGCAGGAC CGGACAGGTG  1500
TATTACAACA CGGATGATGA ACGAGAAGGT GGCAGTGTCC TTGTTAAACG GATGTTCCGG  1560
CCCATGGAAG AGGAGTTTGG TCCGGTGCCT TCAAAGCAGA TGAAAGAAGA AGGGACAAAG  1620
CGAGTGCTCT TGTACGTGAG GAAGGAGACT GACGATGTGT TCGATGCATT GATGTTGAAG  1680
TCTCCCACAG TGATGGGCCT GATGGAAGCG ATATCTGAGA AATATGGGCT GCCCGTGGAG  1740
AAGATAGCAA AGCTTTACAA GAAAAGCAAA AAAGGCATCT TGGTGAACAT GGATGACAAC  1800
ATCATCGAGC ACTACTCGAA CGAGGACACC TTCATCCTCA ACATGGAGAG CATGGTGGAG  1860
GGCTTCAAGG TCACGCTCAT GGAAATCTAG CCCTGGGTTT GGCATCCGCT TTGGCTGGAG  1920
CTCTCAGTGC GTTCCTCCCT GAGAGAGACA GAAGCCCCAG CCCCAGAACC TGGAGACCCA  1980
TCTCCCCCAT CTCACAACTG CTGTTACAAG ACCGTGCTGG GGAGTGGGGC AAGGGACAGG  2040
CCCCACAGTC GGTGTGCTTG GCCCATCCAC TGGCACCTAC CACGGAGCCG AAGCCTGAGC  2100
CCCTCAGGAA GGTGCCTTAG GCCTGTTGGA TTCCTATTTA TTGCCCACCT TTTCCTGGAG  2160
CCCAGGTCCA GGCCCGCCAG GACTCTGCAG GTCACTGCTA GCTCCAGATG AGACCGTCCA  2220
GCGTTCCCCC TTCAAGAGAA ACACTCATCC CGAACAGCCT AAAAAATTCC CATCCCTTCT  2280
TTCTCACCCC TCCATATCTA TATCTCCCGA GTGGCTGGAC AAAATGAGCT ACGTCTGGGT  2340
GCAGTAGTTA TAGGTGGGGC AAGAGGTGGA TGCCCACTTT CTGGTCAGAC ACCTTTAGGT  2400
TGCTCTGGGG AAGGCTGTCT TGCTAAATAC CTCCAGGGTT CCCAGCAAGT GGCCACCAGG  2460
CCTTGTACAG GAAGACATTC AGTCACCGTG TAATTAGTAA CACAGAAAGT CTGCCTGTCT  2520
GCATTGTACA TAGTGTTTAT AATATTGTAA TAATATATTT TACCTGTGGT ATGTGGGCAT  2580
GTTTACTGCC ACTGGCCTAG AGGAGACACA GACCTGGAGA CCGTTTTAAT GGGGGTTTTT  2640
GCCTCTGTGC CTGTTCAAGA GACTTGCAGG GCTAGGTAGA GGGCCTTTGG GATGTTAAGG  2700
TGACTGCAGC TGATGCCAAG ATGGACTCTG CAATGGGCAT ACCTGGGGGC TCGTTCCCTG  2760
TCCCCAGAGG AAGCCCCCTC TCCTTCTCCA TGGGCATGAC TCTCCTTCGA GGCCACCACG  2820
TTTATCTCAC AATGATGTGT TTTGCCTGAC TTTCCCTTTG CGCTGTCTCG TGGGAAAGGT  2880
CATTCTGTCT GAGACCCCAG CTCCTTCTCC AGCTTTGGCT GCGGGCATGG CCTGAGCTTT  2940
CTGGAGAGCC TCTGCAGGGG GTTTGCCATC AGGGCCCTGT GGCTGGGTCT GCTGCAGAGC  3000
TCCTTGGCTA TCAGGAGAAT CCTGGACACT GTACTGTGCC TCCCAGTTTA CAAACACGCC  3060
CTTCATCTCA AGTGGCCCTT TAAAAGGCCT GCTGCCATGT GAGAGCTGTG AACAGCTCAG  3120
CTCTGAGTCG GCAGACTGGG GCTTCCTCCT GGGCCACCAG ATGGAAAGGG GGTATTGTTT  3180
GCCTCACTCC TGGATGCTGC GTTTTAAGGA AGTGAGTGAG AAAGAATGTG CCAAGATACC  3240
TGGCTCCTGT GAAACCAGCC TCAGGAGGGA AACTGGGAGA GAGAAGCTGT GGTCTCCTGC  3300
TACATGCCCT GGGAGCTGGA AGAGAAAAAC ACTCCCCTAA ACAATCGCAA AATGATGAAC  3360
CATCATGGGC CACTGTTCTC TTTGAGGGGA CAGGTTTAGG GGTTTGCGTT CGCCCTTGTG  3420
GGCTGAAGCA CTAGCTTTTT GGTAGCTAGA CACATCCTGC ACCCAAAGGT TCTCTACAAA  3480
GGCCCAGATT TGTTTGTAAA GCACTTTGAC TCTTACCTGG AGGCCCGCTC TCTAAGGGCT  3540
TCCTGCGCTC CCACCTCATC TGTCCCTGAG ATGCAGAGCA GGATGGAAGG TCTGCTTCTA  3600
GCTCAGCTGT TTCTCCTTGA GGTTGCGGAG GAATTGAATT GAATGGGACA GAGGGCAGGT  3660
GCTGTGGCCA AGAAGATCTC CGAGCAGCAG TGACGGGGCA CCTTGCTGTG TGTCCTCTGG  3720
GCATGTTAAC CCTTCTGTGG GGCCAAAGGT TTGCATCGTG GATCCAGCTG TGCTCCAGTC  3780
TGTCCCCTCC TCCTCCACTC TGACTGCCAC GCCCCGGAGC AGCAGCTTGG GGACCCTCCA  3840
GGGTACTAAT GGGGCTCTGT TCTGAGATGG ACAAATTCAG TGTTGGAAAT ACATGTTGTA  3900
CTATGCACTT CCCATGCTCC TAGGGTTAGG AATAGTTTCA AACATGATTG GCAGACATAA  3960
CAACGGCAAA TACTCGGACT GGGGCATAGG ACTCCAGAGT AGGAAAAAGA CAAAAGATTT  4020
GGCAGCCTGA CACAGGCAAC CTACCCCTCT CTCTCCAGCC TCTTTATGAA ACTGTTTGTT  4080
TGCCAGTCCT GCCCTAAGGC AGAAGATGAA TTGAAGATGC TGTGCATGTT TCCTAAGTCC  4140
TTGAGCAATC ATGGTGGTGA CAATTGCCAC AAGGGATATG AGGCCAGTGC CACCAGAGGG  4200
```

```
TGGTGCCAAG TGCCACATCC CTTCCGATCC ATTCCCCTCT GTATCCTCGG AGCACCCCAG 4260
TTTGCCTTTG ATGTGTCCGC TGTGTATGTT AGCTGAACTT TGATGAGCAA AATTTCCTGA 4320
GCGAAACACT CCAAAGAGAT AGGAAAACTT GCCGCCTCTT CTTTTTTGTC CCTTAATCAA 4380
ACTCAAATAA GCTTAAAAAA AATCCATGGA AGATCATGGA CATGTGAAAT GAGCATTTTT 4440
TTCTTTTCTT TTTTTTTTTT TTTTTTTAAC AAAGTCTGAA CTGAACAGAA CAAGACTTTT 4500
TCCTCATACA TCTCCAAATT GTTTAAACTT ACTTTATGAG TGTTTGTTTA GAAGTTCGGA 4560
CCAACAGAAA AATGCAGTCA GATGTCATCT TGGAATTGGT TTCTAAAAGA GTAAGGCATG 4620
TCCCTGCCCA GAAACTTAGG AAGCATGAAA TAAATCAAAT GTTTATTTTC CTTCTTATTT 4680
AAAATCATGC TAATGCAACA GAAATAGAGG GTTTGTGCCA AATGCTATGA ACGGCCCTTT 4740
CTTAAAGACA AGCAAGGGAG ATTGATATAT GTACAATTTG CTCTCATGTT TTT
```

SEQ ID NO:2 BCU4 Protein sequence:
Protein Accession #·   NP_079191.1

```
1    11    21    31    41    51
|    |     |     |     |     |
MSQESDNNKR LVALVPMPSD PPFNTRRAYT SEDEAWKSYL ENPLTAATKA MMIINGDEDS    60
AAALGLLYDY YKVPRDKRLL SVSKASDSQE DQEKRNCLGT SEAQSNLSGG ENRVQVLKTV  120
PVNLSLNQDH LENSKREQYS ISFPESSAII PVSGITVVKA EDFTPVFMAP PVHYPRGDGE  180
EQRVVIFEQT QYDVPSLATH SAYLKDDQRS TPDSTYSESF KDAATEKFRS ASVGAEEYMY  240
DQTSSGTFQY TLEATKSLRQ KQGEGPMTYL NKGQFYAITL SETGDNKCFR HPISKVRSVV  300
MVVFSEDKNR DEQLKYWKYW HSRQHTAKQR VLDIADYKES FNTIGNIEEI AYNAVSFTWD  360
VNEEAKIFIT VNCLSTDFSS QKGVKGLPLM IQIDTYSYNN RSNKPIHRAY CQIKVFCDKG  420
AERKIRDEEQ KQNRKNGKGQ ASQTQCNSSS DGKLAAIPLQ KKSDITYFKT MPDLHSQPVL  480
FIPDVHFANL QRTGQVYYNT DDEREGGSVL VKRMFRPMEE EFGPVPSKQM KEEGTKRVLL  540
YVRKETDDVF DALMLKSPTV MGLMEAISEK YGLPVEKIAK LYKKSKKGIL VNMDDNIIEH  600
YSNEDTFILN MESMVEGFKV TLMEI
```

SEQ ID NO:3 BCU7 DNA SEQUENCE VARIANT 1:
Nucleic Acid Accession #:   AA428062
Coding sequence: 1-777 (entire sequence represents open reading frame)

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGATAGCAA TCTCTGCCGT CAGCAGTGCA CTCCTGTTCT CCCTTCTCTG TGAAGCAAGT    60
ACCGTCGTCC TACTCAATTC CACTGACTCA TCCCCGCCAA CCAATAATTT CACTGATATT   120
GAAGCAGCTC TGAAAGCACA ATTAGATTCA GCGGATATCC CCAAAGCCAG GCGGAAGCGC   180
TACATTTCGC AGAATGACAT GATCGCCATT CTTGATTATC ATAATCAAGT TCGGGGCAAA   240
GTGTTCCCAC CGGCAGCAAA TATGGAATAT ATGGTTTGGG ATGAAAATCT TGCAAAATCG   300
GCAGAGGCTT GGGCGGCTAC TTGCATTTGG GACCATGGAC CTTCTTACTT ACTGAGATTT   360
TTGGGCCAAA ATCTATCTGT ACGCACTGGA AGATATCGCT CTATTCTCCA GTTGGTCAAG   420
CCATGGTATG ATGAAGTGAA AGATTATGCT TTTCCATATC CCCAGGATTG CAACCCCAGA   480
TGTCCTATGA GATGTTTTGG TCCCATGTGC ACACATTATA CGCAGATGGT TTGGGCCACT   540
TCCAATCGGA TAGGATGCGC AATTCATGCT TGCCAAAACA TGAATGTTTG GGGATCTGTG   600
TGGCGACGTG CAGTTTACTT GGTATGCAAC TATGCCCCAA AGGGCAATTG GATTGGAGAA   660
GCACCATATA AAGTAGGGGT ACCATGTTCA TCTTGTCCTC CAAGTTATGG GGGATCTTGT   720
ACTGACAATC TGTGTTTTCC AGGAGTTACG TCAAACTACC TGTACTGGTT TAAA̲TAA̲
```

SEQ ID NO:4 BCU7 DNA SEQUENCE VARIANT 2:
Nucleic Acid Accession #:   AA428062
Coding sequence: 1-777 (entire sequence represents open reading frame)

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGATAGCAA TCTCTGCCGT CAGCAGTGCA CTCCTGTTCT CCCTTCTCTG TGAAGCAAGT    60
ACCGTCGTCC TACTCAATTC CACTGACTCA TCCCCGCCAA CCAATAATTT CACTGATATT   120
GAAGCAGCTC TGAAAGCACA ATTAGATTCA GCGGATATCC CCAAAGCCAG GCGGAAGCGC   180
TACATTTCGC AGAATGACAT GATCGCCATT CTTGATTATC ATAATCAAGT TCGGGGCAAA   240
GTGTTCCCAC CGGCAGCAAA TATGGAATAT ATGGTTTGGG ATGAAAATCT TGCAAAATCG   300
GCAGAGGCTT GGGCGGCTAC TTGCATTTGG GACCATGGAC CTTCTTACTT ACTGAGATTT   360
TTGGGCCAAA ATCTATCTGT ACGCACTGGA AGATATCGCT CTATTCTCCA GTTGGTCAAG   420
CCATGGTATG ATGAAGTGAA AGATTATGCT TTTCCATATC CCCAGGATTG CAACCCCAGA   480
TGTCCTATGA GATGTTTTGG TCCCATGTGC ACACATTATA CGCAGATGGT TTGGGCCACT   540
TCCAATCGGA TAGGATGCGC AATTCATACT TGCCAAAACA TGAATGTTTG GGGATCTGTG   600
TGGCGACGTG CAGTTTACTT GGTATGCAAC TATGCCCCAA AGGGCAATTG GATTGGAGAA   660
GCACCATATA AAGTAGGGGT ACCATGTTCA TCTTGTCCTC CAAGTTATGG GGGATCTTGT   720
ACTGACAATC TGTGTTTTCC AGGAGTTACG TCAAACTACC TGTACTGGTT TAAATAA
```

SEQ ID NO:5 BCU7 Protein sequence Variant 1:
Protein Accession #:   none

```
1          11         21         31         41         51
|          |          |          |          |          |
MIAISAVSSA LLFSLLCEAS TVVLLNSTDS SPPTNNFTDI EAALKAQLDS ADIPKARRKR    60
```

303

```
     YISQNDMIAI LDYHNQVRGK VFPPAANMEY MVWDENLAKS AEAWAATCIW DHGPSYLLRF   120
     LGQNLSVRTG RYRSILQLVK PWYDEVKDYA FPYPQDCNPR CPMRCFGPMC THYTQMVWAT   180
     SNRIGCAIHA CQNMNVWGSV WRRAVYLVCN YAPKGNWIGE APYKVGVPCS SCPPSYGGSC   240
     TDNLCFPGVT SNYLYWFK
```

SEQ ID NO:6 <u>BCU7 Protein sequence Variant 2:</u>
Protein Accession #:   none

```
     1          11         21         31         41         51
     |          |          |          |          |          |
     MIAISAVSSA LLFSLLCEAS TVVLLNSTDS SPPTNNFTDI EAALKAQLDS ADIPKARRKR    60
     YISQNDMIAI LDYHNQVRGK VFPPAANMEY MVWDENLAKS AEAWAATCIW DHGPSYLLRF   120
     LGQNLSVRTG RYRSILQLVK PWYDEVKDYA FPYPQDCNPR CPMRCFGPMC THYTQMVWAT   180
     SNRIGCAIHT CQNMNVWGSV WRRAVYLVCN YAPKGNWIGE APYKVGVPCS SCPPSYGGSC   240
     TDNLCFPGVT SNYLYWFK
```

                              SEQ ID NO:7 BCX2 DNA SEQUENCE
Nucleic Acid Accession #:   NM_003014
Coding sequence: 238-1278 (underlined sequences correspond to start and stop codons)

```
     1   11   21   31   41   51
     |    |    |    |    |    |
     GGCGGGTTCG CGCCCCGAAG GCTGAGAGCT GGCGCTGCTC GTGCCCTGTG TGCCAGACGG    60
     CGGAGCTCCG CGGCCGGACC CCGCGGCCCC GCTTTGCTGC CGACTGGAGT TTGGGGGAAG   120
     AAACTCTCCT GCGCCCCAGA AGATTTCTTC CTCGGCGAAG GGACAGCGAA AGATGAGGGT   180
     GGCAGGAAGA GAAGGCGCTT TCTGTCTGCC GGGGTCGGCG CGCGAGAGGG CAGTGCCATG   240
     TTCCTCTCCA TCCTAGTGGC GCTGTGCCTG TGGCTGCACC TGGCGCTGGG CGTGCGCGGC   300
     GCGCCCTGCG AGGCGGTGCG CATCCCTATG TGCCGGCACA TGCCCTGGAA CATCACGCGG   360
     ATGCCCAACC ACCTGCACCA CAGCACGCAG GAGAACGCCA TCCTGGCCAT CGAGCAGTAC   420
     GAGGAGCTGG TGGACGTGAA CTGCAGCGCC GTGCTGCGCT TCTTCTTCTG TGCCATGTAC   480
     GCGCCCATTT GCACCCTGGA GTTCCTGCAC GACCCTATCA AGCCGTGCAA GTCGGTGTGC   540
     CAACGCGCGC GCGACGACTG CGAGCCCCTC ATGAAGATGT ACAACCACAG CTGGCCCGAA   600
     AGCCTGGCCT GCGACGAGCT GCCTGTCTAT GACCGTGGCG TGTGCATTTC GCCTGAAGCC   660
     ATCGTCACGG ACCTCCCGGA GGATGTTAAG TGGATAGACA TCACACCAGA CATGATGGTA   720
     CAGGAAAGGC CTCTTGATGT TGACTGTAAA CGCCTAAGCC CCGATCGGTG CAAGTGTAAA   780
     AAGGTGAAGC CAACTTTGGC AACGTATCTC AGCAAAAACT ACAGCTATGT TATTCATGCC   840
     AAAATAAAAG CTGTGCAGAG GAGTGGCTGC AATGAGGTCA CAACGGTGGT GGATGTAAAA   900
     GAGATCTTCA AGTCCTCATC ACCCATCCCT CGAACTCAAG TCCCGCTCAT TACAAATTCT   960
     TCTTGCCAGT GTCCACACAT CCTGCCCCAT CAAGATGTTC TCATCATGTG TTACGAGTGG  1020
     CGTTCAAGGA TGATGCTTCT TGAAAATTGC TTAGTTGAAA AATGGAGAGA TCAGCTTAGT  1080
     AAAAGATCCA TACAGTGGGA AGAGAGGCTG CAGGAACAGC GGAGAACAGT TCAGGACAAG  1140
     AAGAAAACAG CCGGGCGCAC CAGTCGTAGT AATCCCCCCA AACCAAAGGG AAAGCCTCCT  1200
     GCTCCCAAAC CAGCCAGTCC CAAGAAGAAC ATTAAAACTA GGAGTGCCCA GAAGAGAACA  1260
     AACCCGAAAA GAGTGTGAGC TAACTAGTTT CCAAAGCGGA GACTTCCGAC TTCCTTACAG  1320
     GATGAGGCTG GGCATTGCCT GGGACAGCCT ATGTAAGGCC ATGTGCCCCT TGCCCTAACA  1380
     ACTCACTGCA GTGCTCTTCA TAGACACATC TTGCAGCATT TTTCTTAAGG CTATGCTTCA  1440
     GTTTTTCTTT GTAAGCCATC ACAAGCCATA GTGGTAGGTT TGCCCTTTGG TACAGAAGGT  1500
     GAGTTAAAGC TGGTGGAAAA GGCTTATTGC ATTGCATTCA GGATGAACCTG TGTGCATACT  1560
     CTAGAAGAGT AGGGGAAAATA ATGCTTGTTA CAATTCGACC TAATATGTGC ATTGTAAAAT  1620
     AAATGCCATA TTTCAAACAA AACACGTAAT TTTTTTTACAG TATGTTTTAT TACCTTTTGA  1680
     TATCTGTTGT TGCAATGTTA GTGATGTTTT AAAATGTGAT GAAAATAAA TGTTTTTAAG  1740
     AAGGAACAGT AGTGGAATGA ATGTTAAAAG ATCTTTATGT GTTTATGGTC TGCAGAAGGA  1800
     TTTTTGTGAT GAAAGGGGAT TTTTTGAAAA ATTAGAGAAG TAGCATATGG AAAATTATAA  1860
     TGTGTTTTTT TACCAATGAC TTCAGTTTCT GTTTTTAGCT AGAAACTTAA AAACAAAAAT  1920
     AATAATAAAG AAAAATAAAT AAAAAGGAGA GGCAGAACAAT GTCTGGATTC CTGTTTTTTG  1980
     GTTACCTGAT TTCCATGATC ATGATGCTTC TTGTCAACAC CCTCTTAAGC AGCACCAGAA  2040
     ACAGTGAGTT TGTCTGTACC ATTAGGAGTT AGGTACTAAT TAGTTGGCTA ATGCTCAAGT  2100
     ATTTTATACC CACAAGAGAG GTATGTCACT CATCTTACTT CCCAGGACAT CCACCCTGAG  2160
     AATAATTTGA CAAGCTTAAA AATGGCCTTC ATGTGAGTGC CAAATTTTGT TTTTCTTCAT  2220
     TTAAATATTT TCTTTGCCTA AATACATGTG AGAGGAGTTA AATATAAATG TACAGAGAGG  2280
     AAAGTTGAGT TCCACCTCTG AAATGAGAAT TACTTGACAG TTGGGATACT TTAATCAGAA  2340
     AAAAAGAACT TATTTGCAGC ATTTTATCAA CAAATTTCAT AATTGTGGAC AATTGGAGGC  2400
     ATTTATTTTA AAAAACAATT TTATTGGCCT TTTGCTAACA CAGTAAGCAT GTATTTTATA  2460
     AGGCATTCAA TAAATGCACA ACGCCCAAAG GAAATAAAAT CCTATCTAAT CCTACTCTCC  2520
     ACTACACAGA GGTAATCACT ATTAGTATTT TGGCATATTA TTCTCCAGGT GTTTGCTTAT  2580
     GCACTTATAA AATGATTTGA ACAAATAAAA CTAGGAACCT GTATACATGT GTTTCATAAC  2640
     CTGCCTCCTT TGCTTGGCCC TTTATTGAGA TAAGTTTTCC TGTCAAGAAA GCAGAAACCA  2700
     TCTCATTTCT AACAGCTGTG TTATATTCCA TAGTATGCAT TACTCAACAA ACTGTTGTGC  2760
     TATTGGATAC TTAGGTGGTT TCTTCACTGA CAATACTGAA TAAACATCTC ACCGGAATTC
```

SEQ ID NO:8 <u>BCX2 Protein sequence:</u>
Protein Accession #: NP_003005.1

```
     1    11   21   31   41   51
     |    |    |    |    |    |
     MFLSILVALC LWLHLALGVR GAPCEAVRIP MCRHMPWNIT RMPNHLHHST QENAILAIEQ   60
```

```
      YEELVDVNCS AVLRFFFCAM YAPICTLEFL HDPIKPCKSV CQRARDDCEP LMKMYNHSWP  120
      ESLACDELPV YDRGVCISPE AIVTDLPEDV KWIDITPDMM VQERPLDVDC KRLSPDRCKC  180
      KKVKPTLATY LSKNYSYVIH AKIKAVQRSG CNEVTTVVDV KEIFKSSSPI PRTQVPLITN  240
      SSCQCPHILP HQDVLIMCYE WRSRMMLLEN CLVEKWRDQL SKRSIQWEER LQEQRRTVQD  300
5     KKKTAGRTSR SNPPKPKGKP PAPKPASPKK NIKTRSAQKR TNPKRV
```

SEQ ID NO:9 CBK1 DNA SEQUENCE

Nucleic Acid Accession #:   NM_032391

10 Coding sequence: 129-302 (underlined sequences correspond to start and stop codons)

```
      1          11         21         31         41         51
      |          |          |          |          |          |
15    GTCCTTCCTC TCCTAGCCTA AGGCGTGCAA ACAGAGCGCC ACTGGGAGGC TGAAACCTTT    60
      AGGCCGATGC TTGCTTGCAA GGTCAGGCAA GCTGGATTCT GGTCCCCACC TTTGCAGAGA   120
      GAACAGCGAT GTTGTGCGCC CATTTCTCAG ATCAAGGACC GGCCCATCTT ACTACCTCCA   180
      AGAGTGCTTT TCTCTCTAAT AAGAAAACAT CTACTTTGAA ACATCTACTG GGCGAGACCA   240
      GGAGTGATGG CTCAGCCTGT AATTCTGGAA TTTCGGGAGG CCGAGGCAGG AAGATTCCTT   300
20    GAGCACAGGA GTTCCAGACC AGCCTGGGCA ATGTAGCAAG ACGCTGTCTC TATTTATACA   360
      ATAAAATTTT TTTAAAAAAG G
```

25 **SEQ ID NO:10** CBK1 Protein sequence:

Protein Accession #:   NP_115767

```
      1          11         21         31         41         51
      |          |          |          |          |          |
30    MLCAHFSDQG PAHLTTSKSA FLSNKKTSTL KHLLGETRSD GSACNSGISG GRGRKIP
```

SEQ ID NO:11 CHA1 DNA SEQUENCE

Nucleic Acid Accession #:   NM_020182

35 Coding sequence: 96-854 (underlined sequences correspond to start and stop codons)

```
      1          11         21         31         41         51
      |          |          |          |          |          |
40    TCCTTGGGTT CGGGTGAAAG CGCCTGGGGG TTCGTGGCCA TGATCCCCGA GCTGCTGGAG    60
      AACTGAAGGC GGACAGTCTC CTGCGAAACC AGGCAATGGA GGAGCTGGAG TTTGTTCAGA   120
      TCATCATCAT CGTGGTGGTG ATGATGGTGA TGGTGGTGGT GATCACGTGC CTGCTGAGCC   180
      ACTACAAGCT GTCTGCACGG TCCTTCATCA GCCGGCACAC CCAGGGGCGG AGGAGAGAAG   240
      ATGCCCTGTC CTCAGAAGGA TGCCTGTGGC CCTCGGAGAG CACAGTGTCA GGCAACGGAA   300
45    TCCCAGAGCG GCAGGTCTAC GCCCCGCCTC GGCCCACCGA CCGCCTGGCC GTGCCGCCCT   360
      TCGCCCAGCG GGAGCGCTTC CACCGCTTCC AGCCCACCTA TCCGTACCTG CAGCACGAGA   420
      TCGACCTGCC ACCCACCATC TCGCTGTCAG ACGGGGAGGA GCCCCCACCC TACCAGGGCC   480
      CCTGCACCCT CCAGCTTCGG GACCCCGAGC AGCAGCTGGA ACTGAACCGG GAGTCGGTGC   540
      GCGCACCCCC AAACAGAACC ATCTTCGACA GTGACCTGAT GGATAGTGCC AGGCTGGGCG   600
50    GCCCCTGCCC CCCCAGCAGT AACTCGGGCA TCAGCGCCAC GTGCTACGGC AGCGGCGGGC   660
      GCATGGAGGG GCCGCCGCCC ACCTACAGCG AGGTCATCGG CCACTACCCG GGGTCCTCCT   720
      TCCAGCACCA GCAGAGCAGT GGGCCGCCCT CCTTGCTGGA GGGGACCCGG CTCCACCACA   780
      CACACATCGC GCCCCTAGAG AGCGCAGCCA TCTGGACAA AGAGAAGGAT AAACAGAAAG   840
      GACACCCTCT CTAGGGTCCC CAGGGGGGCC GGGCTGGGGC TGCGTAGGTG AAAAGGCAGA   900
55    ACACTCCGCG CTTCTTAGAA GAGGAGTGAG AGGAAGGCGG GGGGCGCAGC AACGCATCGT   960
      GTGGCCCTCC CCTCCCACCT CCCTGTGTAT AAATATTTAC ATGTGATGTC TGGTCTGAAT  1020
      GCACAAGCTA AGAGAGCTTG CAAAAAAAAA AAGAAAAAAG AAAAAAAAAA ACCACGTTTC  1080
      TTTGTTGAGC TGTGTCTTGA AGGCAAAAGA AAAAAAATTT CTACAGTAAA AAAAAAAAAA  1140
60    A
```

SEQ ID NO:12 CHA1 Protein sequence:

Protein Accession #:   NP_064567

```
      1          11         21         31         41         51
65    |          |          |          |          |          |
      MAELEFVQII IIVVVMMVMV VVITCLLSHY KLSARSFISR HSQGRRREDA LSSEGCLWPS    60
      ESTVSGNGIP EPQVYAPPRP TDRLAVPPFA QRERFHRFQP TYPYLQHEID LPPTISLSDG   120
      EEPPPYQGPC TLQLRDPEQQ LELNRESVRA PPNRTIFDSD LMDSARLGGP CPPSSNSGIS   180
70    ATCYGSGGRM EGPPPTYSEV IGHYPGSSFQ HQQSSGPPSL LEGTRLHHTH IAPLESAAIW   240
      SKEKDKQKGH PL
```

SEQ ID NO:13 CJA5 DNA SEQUENCE

75 Nucleic Acid Accession #:   NM_012445

Coding sequence: 276-1271 (underlined sequences correspond to start and stop codons)

```
      1          11         21         31         41         51
80    |          |          |          |          |          |
      |          |          |          |          |          |
```

305

```
GCACGAGGGA AGAGGGTGAT CCGACCCGGG GAAGGTCGCT GGGCAGGGCG AGTTGGGAAA    60
GCGGCAGCCC CCGCCGCCCC CGCAGCCCCT TCTCCTCCTT TCTCCCACGT CCTATCTGCC   120
TCTCGCTGGA GGCCAGGCCG TGCAGCATCG AAGACAGGAG GAACTGGAGC CTCATTGGCC   180
GGCCCGGGGC GCCGGCCTCG GGCTTAAATA GGAGCTCCGG GCTCTGGCTG GGACCCGACC   240
GCTGCCGGCC GCGCTCCCGC TGCTCCTGCC GGGTGATGGA AAACCCCAGC CCGGCCGCCG   300
CCCTGGGCAA GGCCCTCTGC GCTCTCCTCC TGGCCACTCT CGGCGCCGCC GGCCAGCCTC   360
TTGGGGGAGA GTCCATCTGT TCCGCCAGAG CCCCCGGCCAA ATACAGCATC ACCTTCACGG   420
GCAAGTGGAG CCAGACGGCC TTCCCCAAGC AGTACCCCCT GTTCCGCCCC CCTGCGCAGT   480
GGTCTTCGCT GCTGGGGGCC GCGCATAGCT CCGACTACAG CATGTGGAGG AAGAACCAGT   540
ACGTCAGTAA CGGGCTGCGC GACTTTGCGG AGCGCGGCGA GGCCTGGGCG CTGATGAAGG   600
AGATCGAGGC GGCGGGGGAG GCGCTGCAGA GCGTGCACGC GGTGTTTTCG GCGCCCGCCG   660
TCCCCAGCGG CACCGGGCAG ACGTCGGCGG AGCTGGAGGT GCAGCGCAGG CACTCGCTGG   720
TCTCGTTTGT GGTGCGCATC GTGCCCAGCC CCGACTGGTT CGTGGGCGTG GACAGCCTGG   780
ACCTGTGCGA CGGGGACCGT TGGCGGGAAC AGGCGGCGCT GGACCTGTAC CCCTACGACG   840
CCGGGACGGA CAGCGGCTTC ACCTTCTCCT CCCCCAACTT CGCCACCATC CCGCAGGACA   900
CGGTGACCGA GATAACGTCC TCCTCTCCCA GCCACCCGGC CAACTCCTTC TACTACCCGC   960
GGCTGAAGGC CCTGCCTCCC ATCGCCAGGG TGACACTGGT GCGGCTGCGA CAGAGCCCCA  1020
GGGCCTTCAT CCCTCCCGCC CCAGTCCTGC CCAGCAGGGA CAATGAGATT GTAGACAGCG  1080
CCTCAGTTCC AGAAACGCCG CTGGACTGCG AGGTCTCCCT GTGGTCGTCC TGGGGACTGT  1140
GCGGGAGGCCA CTGTGGGGAGG CTCGGGACCA AGAGCAGGAC TCGCTACGTC CGGGTCCAGC  1200
CCGCCAACAA CGGGGAGCCCC TGCCCCGAGC TCGAAGAAGA GGCTGAGTGC GTCCCTGATA  1260
ACTGCGTCTA AGACCAGAGC CCCGCAGCCC CTGGGGCCCC CGGAGCCATG GGGTGTCGGG  1320
GGCTCCTGTG CAGGCTCATG CTGCAGGCGG CCGAGGCACA GGGGGTTTCG CGCTGCTCCT  1380
GACCGCGGTG AGGCCGCGCC GACCATCTCT GCACTGAAGG GCCCTCTGGT GGCCGGCACG  1440
GGCATTGGGA AACAGCCTCC TCCTTTCCCA ACCTTGCTTC TTAGGGGCCC CCGTGTCCCG  1500
TCTGCTCTCA GCCTCCTCCT CCTGCAGGAT AAAGTCATCC CCAAGGCTCC AGCTACTCTA  1560
AATTATGGTC TCCTTATAAG TTATTGCTGC TCCAGGAGAT TGTCCTTCAT CGTCCAGGGG  1620
CCTGGCTCCC ACGTGGTTGC AGATACCTCA GACCTGGTGC TCTAGGCTGT GCTGAGCCCA  1680
CTCTCCCGAG GGCGCATCCA AGCGGGGGCC ACTTGAGAAG TGAATAAATG GGGCGGTTTC  1740
GGAAGCGTCA GTGTTTCCAT GTTATGGATC TCTCTGCGTT TGAATAAAGA CTATCTCTGT  1800
TGCTCAC
```

SEQ ID NO:14 CJA5 Protein sequence:
Protein Accession #:   NP_036577

```
1          11         21         31         41         51
|          |          |          |          |          |
MENPSPAAAL GKALCALLLA TLGAAGQPLG GESICSARAP AKYSITFTGK WSQTAFPKQY    60
PLFRPPAQWS SLLGAAHSSD YSMWRKNQYV SNGLRDFAER GEAWALMKEI EAAGEALQSV   120
HAVFSAPAVP SGTGQTSAEL EVQRRHSLVS FVVRIVPSPD WFVGVDSLDL CDGDRWREQA   180
ALDLYPYDAG TDSGFTFSSP NFATIPQDTV TEITSSSPSH PANSFYYPRL KALPPIARVT   240
LVRLRQSPRA FIPPAPVLPS RDNEIVDSAS VPETPLDCEV SLWSSWGLCG GHCGRLGTKS   300
RTRYVRVQPA NNGSPCPELE EEAECVPDNC V
```

SEQ ID NO:15 LBH9 DNA SEQUENCE

Nucleic Acid Accession #:   NM_002391
Coding sequence: 26-457 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
CGGGCGAAGC AGCGCGGGCA GCGAGATGCA GCACCGAGGC TTCCTCCTCC TCACCCTCCT    60
CGCCCTGCTG GCGCTCACCT CCGCGGTCGC CAAAAAGAAA GATAAGGTGA AGAAGGGCGG   120
CCCCGGGGAGC GAGTGCGCTG AGTGGGCCTG GGGGCCCTGC ACCCCCAGCA GCAAGGATTG   180
CGGCGTGGGT TTCCGCGAGG GCACCTGCGG GGCCCAGACC CAGCGCATCC GGTGCAGGGT   240
GCCCTGCAAC TGGAAGAAGG AGTTTGGAGC CGACTGCAAG TACAAGTTTG AGAACTGGGG   300
TGCGTGTGAT GGGGGCACAG GCACCAAAGT CCGCCAAGGC ACCCTGAAGA AGGCGCGCTA   360
CAATGCTCAG TGCCAGGAGA CCATCCGCGT CACCAAGCCC TGCACCCCCA AGACCAAAGC   420
AAAGGCCAAA GCCAAGAAAG GGAAGGGAAA GGACTAGACG CCAAGCCTGG ATGCCAAGGA   480
GCCCCTGGTG TCACATGGGG CCTGGCCACG CCCTCCCTCT CCCAGGCCCG AGATGTGACC   540
CACCAGTGCC TTCTGTCTGC TCGTTAGCTT TAATCAATCA TGCCCTGCCT TGTCCCTCTC   600
ACTCCCCAGC CCCACCCCTA AGTGCCCAAA GTGGGGAGGG ACAAGGGATT CTGGGAAGCT   660
TGAGCCTCCC CCAAAGCAAT GTGAGTCCCA GAGCCCGCTT TTGTTCTTCC CCACAATTCC   720
ATTACTAAGA AACACATCAA ATAAACTGAC TTTTTCCCCC CAATAAAAGC TCTTCTTTTT   780
TAAATAT
```

SEQ ID NO:16 LBH9 Protein sequence:
Protein Accession #:   NP_002382

```
1          11         21         31         41         51
|          |          |          |          |          |
MQHRGFLLLT LLALLALTSA VAKKKDKVKK GGPGSECAEW AWGPCTPSSK DCGVGFREGT    60
CGAQTQRIRC RVPCNWKKEF GADCKYKFEN WGACDGGTGT KVRQGTLKKA RYNAQCQETI   120
RVTKPCTPKT KAKAKAKKGK GKD
```

<div align="center">SEQ ID NO:17 LEM9 DNA SEQUENCE</div>

Nucleic Acid Accession #:   NM_005244
Coding sequence: 1-1617 (underlined sequences correspond to start and stop codons)

```
        1          11         21         31         41         51
        |          |          |          |          |          |
     ATGGTAGAAC TAGTGATCTC ACCCAGCCTC ACTGTAAACA GCGATTGTCT GGATAAACTG   60
     AAGTTTAACC GTGCTGACGC TGCTGTGTGG ACTCTGAGTG ACAGACAAGG CATCACCAAA  120
     TCGGCCCCCC TGAGAGTGTC CCAGCTCTTC TCCAGATCTT GCCCACGTGT CCTCCCCCGC  180
     CAGCCTTCCA CAGCCATGGC AGCCTACGGC CAGACGCAGT ACAGTGCGGG GATCCAGCAG  240
     GCTACCCCCT ATACAGCTTA CCCACCTCCA GCACAAGCCT ATGGAATCCC TTCCTACAGC  300
     ATCAAGACAG AAGACAGCTT GAACCATTCC CCTGGCCAGA GTGGATTCCT CAGCTATGGC  360
     TCCAGCTTCA GCACCTCACC CACTGGACAG AGCCCATACA CCTACCAGAT GCACGGCACA  420
     ACAGGGTTCT ATCAAGGAGG AAATGGACTG GGCAACGCAG CCGGTTTCGG GAGTGTGCAC  480
     CAGGACTATC CTTCCTACCC CGGCTTCCCC CAGAGCCAGT ACCCCCAGTA TTACGGCTCA  540
     TCCTACAACC CTCCCTACGT CCCGGCCAGC AGCATCTGCC CTTCGCCCCT CTCCACGTCC  600
     ACCTACGTCC TCCAGGAGGC ATCTCACAAC GTCCCCAACC AGAGTTCCGA GTCACTTGCT  660
     GGTGAATACA ACACACACAA TGGACCTTCC ACACCAGCGA AAGAGGGAGA CACAGACAGG  720
     CCGCACCGGG CCTCCGACGG GAAGCTCCGA GGCCGGTCTA AGAGGAGCAG TGACCCGTCC  780
     CCGGCAGGGG ACAATGAGAT TGAGCGTGTG TTCGTGTGGG ACTTGGATGA GACAATAATT  840
     ATTTTTCACT CCTTACTCAC GGGGACATTT GCATCCGAT ACGGGAAGGA CACCACGACG  900
     TCCGTGCGCA TTGGCCTTAT GATGGAAGAA ATGATCTTCA ACCTTGCAGA TACACATCTG  960
     TTCTTCAATG ACCTGGAGGA TTGTGACCAG ATCCACGTTG ATGACGTCTC ATCAGATGAC 1020
     AATGGCCAAG ATTTAAGCAC ATACAACTTC TCCGCTGACG GCTTCCACAG TTCGGCCCCA 1080
     GGAGCCAACC TGTGCCTGGG CTCTGGCGTG CACGGCGGCG TGGACTGGAT GAGGAAGCTG 1140
     GCCTTCCGCT ACCGGCGGGT GAAGGAGATG TACAATACCT ACAAGAACAA CGTTGGTGGG 1200
     TTGATAGGCA CTCCCAAAAG GGAGACCTGG CTACAGCTCC GAGCTGAGCT GGAAGCTCTC 1260
     ACAGACCTCT GGCTGACCCA CTCCCTGAAG GCACTAAACC TCATCAACTC CCGGCCCAAC 1320
     TGTGTCAATG TGCTGGTCAC CACCACTCAA CTAATTCCTG CCCTGGCCAA AGTCCTGCTA 1380
     TATGGCCTGG GGTCTGTGTT TCCTATTGAG AACATCTACA GTGCAACCAA GACAGGGAAG 1440
     GAGAGCTGCT TCGAGAGGAT AATGCAGAGA TTCGGCAGAA AAGCTGTCTA CGTGGTGATC 1500
     GGTGATGGTG TGGAAGAGGA GCAAGGAGCG AAAAAGCACA ACATGCCTTT CTGGCGGATA 1560
     TCCTGCCACG CAGACCTGGA GGCACTGAGG CACGCCCTGG AACTGGAGTA TTTATAG
```

SEQ ID NO:18 LEM9 Protein sequence:
Protein Accession #:   NP_005235

```
        1          11         21         31         41         51
        |          |          |          |          |          |
     MVELVISPSL TVNSDCLDKL KFNRADAAVW TLSDRQGITK SAPLRVSQLF SRSCPRVLPR   60
     QPSTAMAAYG QTQYSAGIQQ ATPYTAYPPP AQAYGIPSYS IKTEDSLNHS PGQSGFLSYG  120
     SSFSTSPTGQ SPYTYQMHGT TGFYQGGNGL GNAAGFGSVH QDYPSYPGFP QSQYPQYYGS  180
     SYNPPYVPAS SICPSPLSTS TYVLQEASHN VPNQSSESLA GEYNTHNGPS TPAKEGDTDR  240
     PHRASDGKLR GRSKRSSDPS PAGDNEIERV FVWDLDETII IFHSLLTGTF ASRYGKDTTT  300
     SVRIGLMMEE MIFNLADTHL FFNDLEDCDQ IHVDDVSSDD NGQDLSTYNF SADGFHSSAP  360
     GANLCLGSGV HGGVDWMRKL AFRYRRVKEM YNTYKNNVGG LIGTPKRETW LQLRAELEAL  420
     TDLWLTHSLK ALNLINSRPN CVNVLVTTTQ LIPALAKVLL YGLGSVFPIE NIYSATKTGK  480
     ESCFERIMQR FGRKAVYVVI GDGVEEEQGA KKHNMPFWRI SCHADLEALR HALELEYL
```

<div align="center">SEQ ID NO:19 OAA1 DNA SEQUENCE</div>

Nucleic Acid Accession #:   NM_002740
Coding sequence: 178-1968 (underlined sequences correspond to start and stop codons)

```
        1          11         21         31         41         51
        |          |          |          |          |          |
     CCGCGGTTCC GGCTGCTCCG GCGAGGCGAC CCTTGGGTCG GCGCTGCGGG CGAGGTGGGC   60
     AGGTAGGTGG GCGGACGGCC GCGGTTCTCC GGCAAGCGCA GGCGGCGGAG TCCCCCACGG  120
     CGCCCGAAGC GCCCCCCGCA CCCCCGGCCT CCAGCGTTGA GGCGGGGGAG TGAGGAGATG  180
     CCGACCCAGA GGGACAGCAG CACCATGTCC CACACGGTCG CAGGCGGCGG CAGCGGGGAC  240
     CATTCCCACC AGGTCCGGGT GAAAGCCTAC TACCGCGGGG ATATCATGAT AACACATTTT  300
     GAACCTTCCA TCTCCTTTGA GGGCCTTTGC AATGAGGTTC GAGACATGTG TTCTTTTGAC  360
     AACGAACAGC TCTTCACCAT GAAATGGATA GATGAGGAAG GAGACCCGTG TACAGTATCA  420
     TCTCAGTTGG AGTTAGAAGA AGCCTTTAGA CTTTATGAGC TAAACAAGGA TTCTGAACTC  480
     TTGATTCATG TGTTCCCTTG TGTACCAGAA CGTCCTGGGA TGCCTTGTCC AGGAGAAGAT  540
     AAATCCATCT ACCGTAGAGG TGCACGCCGC TGGAGAAAGC TTTATTGTGC CAATGGCCAC  600
     ACTTTCCAAG CCAAGCGTTT CAACAGGCGT GCTCACTGTG CCATCTGCAC AGACCGAATA  660
     TGGGGACTTG GACGCCAAGG ATATAAGTGC ATCAACTGCA AACTCTTGGT TCATAAGAAG  720
     TGCCATAAAC TCGTCACAAT TGAATGTGGG CGGCATTCTT TGCCACAGGA ACCAGTGATG  780
     CCCATGGATC AGTCATCCAT GCATTCTGAC CATGCACAGA CAGTAATTCC ATATAATCCT  840
     TCAAGTCATG AGAGTTTGGA TCAAGTTGGT GAAGAAAAAG AGGCAATGAA CACCAGGGAA  900
     AGTGGCAAAG CTTCATCCAG TCTAGGTCTT CAGGATTTTG ATTTGCTCCG GGTAATAGGA  960
     AGAGGAAGTT ATGCCAAAGT ACTGTTGGTT CGATTAAAAA AAACAGATCG TATTTATGCA 1020
     ATGAAAGTTG TGAAAAAGA GCTTGTTAAT GATGATGAGG ATATTGATTG GGTACAGACA 1080
     GAGAAGCATG TGTTTGAGCA GGCATCCAAT CATCCTTTCC TTGTTGGGCT GCATTCTTGC 1140
     TTTCAGACAG AAAGCAGATT GTTCTTTGTT ATAGAGTATG TAAATGGAGG AGACCTAATG 1200
     TTTCATATGC AGCGACAAAG AAAACTTCCT GAAGAACATG CCAGATTTTA CTCTGCAGAA 1260
```

```
ATCAGTCTAG CATTAAATTA TCTTCATGAG CGAGGGATAA TTTATAGAGA TTTGAAACTG   1320
GACAATGTAT TACTGGACTC TGAAGGCCAC ATTAAACTCA CTGACTACGG CATGTGTAAG   1380
GAAGGATTAC GGCCAGGAGA TACAACCAGC ACTTTCTGTG GTACTCCTAA TTACATTGCT   1440
CCTGAAATTT TAAGAGGAGA AGATTATGGT TTCAGTGTTG ACTGGTGGGC TCTTGGAGTG   1500
CTCATGTTTG AGATGATGGC AGGAAGGTCT CCATTTGATA TTGTTGGGAG CTCCGATAAC   1560
CCTGACCAGA ACACAGAGGA TTATCTCTTC CAAGTTATTT TGGAAAAACA AATTCGCATA   1620
CCACGTTCTC TGTCTGTAAA AGCTGCAAGT GTTCTGAAGA GTTTTCTTAA TAAGGACCCT   1680
AAGGAACGAT TGGGTTGTCA TCCTCAAACA GGATTTGCTG ATATTCAGGG ACACCCGTTC   1740
TTCCGAAATG TTGATTGGGA TATGATGGAG CAAAAACAGG TGGTACCTCC CTTTAAACCA   1800
AATATTTCTG GGGAATTTGG TTTGGACAAC TTTGATTCTC AGTTTACTAA TGAACCTGTC   1860
CAGCTCACTC CAGATGACGA TGACATTGTG AGGAAGATTG ATCAGTCTGA ATTTGAAGGT   1920
TTTGAGTATA TCAATCCTCT TTTGATGTCT GCAGAAGAAT GTGTCTGATC CTCATTTTTC   1980
AACCATGTAT TCTACTCATG TTGCCATTTA ATGCATGGAT AAACTTGCTG CAAGCCTGGA   2040
TACAATTAAC CATTTTATAT TTGCCACCTA CAAAAAAACA CCCAATATCT TCTCTTGTAG   2100
ACTATATGAA TCAATTATTA CATCTGTTTT ACTATGAAAA AAAAATTAAT ACTACTAGCT   2160
TCCAGACAAT CATGTCAAAA TTTAGTTGAA CTGGTTTTTC AGTTTTTAAA AGGCCTACAG   2220
ATGAGTAATG AAGTTACCTT TTTTGTTTAA AAAAAAAAAA G
```

```
 1          11         21         31         41         51
 |          |          |          |          |          |
MSHTVAGGGS GDHSHQVRVK AYYRGDIMIT HFEPSISFEG LCNEVRDMCS FDNEQLFTMK    60
WIDEEGDPCT VSSQLELEEA FRLYELNKDS ELLIHVFPCV PERPGMPCPG EDKSIYRRGA   120
RRWRKLYCAN GHTFQAKRFN RRAHCAICTD RIWGLGRQGY KCINCKLLVH KKCHKLVTIE   180
CGRHSLPQEP VMPMDQSSMH SDHAQTVIPY NPSSHESLDQ VGEEKEAMNT RESGKASSSL   240
GLQDFDLLRV IGRGSYAKVL LVRLKKTDRI YAMKVVKKEL VNDDEDIDWV QTEKHVFEQA   300
SNHPFLVGLH SCFQTESRLF FVIEYVNGGD LMFHMQRQRK LPEEHARFYS AEISLALNYL   360
HERGIIYRDL KLDNVLLDSE GHIKLTDYGM CKEGLRPGDT TSTFCGTPNY IAPEILRGED   420
YGFSVDWWAL GVLMFEMMAG RSPFDIVGSS DNPDQNTEDY LFQVILEKQI RIPRSLSVKA   480
ASVLKSFLNK DPKERLGCHP QTGFADIQGH PFFRNVDWDM MEQKQVVPPF KPNISGEFGL   540
DNFDSQFTNE PVQLTPDDDD IVRKIDQSEF EGFEYINPLL MSAEECV
```

```
 1          11         21         31         41         51
 |          |          |          |          |          |
CCAGGCGGCG TTGCGGCCCC GGCCCCGGCT CCCTGCGCCG CCGCCGCCGC CGCCGCCGCC    60
GCCGCCGCCG CCGCCGCCAG CGCTAGCGCC AGCAGCCGGG CCCGATCACC CGCCGCCCGG   120
TGCCCGCCGC CGCCCGCGCC AGCAACCGGG CCCGATCACC CGCCGCCCGG TGCCCGCCGC   180
CGCCCGCCGCC ACCGGCATGG CGCTCCGGGG CTTCTGCAGC GCCGATGGCT CCGACCCGCT   240
CTGGGACTGG AATGTCACGT GGAATACCAG CAACCCCGAC TTCACCAAGT GCTTTCAGAA   300
CACGGTCCTC GTGTGGGTGC CTTGTTTTTA CCTCTGGGCC TGTTTCCCCT TCTACTTCCT   360
CTATCTCTCC CGACATGACC GAGGCTACAT TCAGATGACA CCTCTCAACA AAACCAAAAC   420
TGCCTTGGGA TTTTTGCTGT GGATCGTCTG CTGGGCAGAC CTCTTCTACT CTTTCTGGGA   480
AAGAAGTCGG GGCATATTCC TGGCCCCAGT GTTTCTGGTC AGCCCAACTC TCTTGGGCAT   540
CACCACGCTG CTTGCTACCT TTTTAATTCA GCTGGAGAGG AGGAAGGGAG TTCAGTCTTC   600
AGGGATCATG CTCACTTTCT GGCTGGTAGC CCTAGTGTGT GCCCTAGCCA TCCTGAGATC   660
CAAAATTATG ACAGCCTTAA AAGAGGATGC CCAGGTGGAC CTGTTTCGTG ACATCACTTT   720
CTACGTCTAC TTTTCCCTCT TACTCATTCA GCTCGTCTTG TCCTGTTTCT CAGATCGCTC   780
ACCCCTGTTC TCGGAAACCA TCCACGACCC TAATCCCTGC CCAGAGTCCA GCGCTTCCTT   840
CCTGTCGAGG ATCACCTTCT GGTGGATCAC AGGGTTGATT GTCCGGGGCT ACCGCCAGCC   900
CCTGGAGGGC AGTGACCTCT GGTCCTTAAA CAAGGAGGAC ACGTCGGAAC AAGTCGTGCC   960
TGTTTTGGTA AAGAACTGGA AGAAGGAATG CGCCAAGACT AGGAAGCAGC CGGTGAAGGT  1020
TGTGTACTCC TCCAAGGATC CTGCCCAGCC GAAAGATGAT TCCAAGGTGG ATGCGAATGA  1080
GGAGGTGGAG GCTTTGATCG TCAAGTCCCC ACAGAAGGAG TGGAACCCCT CTCTGTTTAA  1140
GGTGTTATAC AAGACCTTTG GGCCCTACTT CCTCATGAGC TTCTTCTTCA AGGCCATCCA  1200
CGACCTGATG ATGTTTTCCG GGCCGCAGAT CTTAAAGTTG CTCATCAAGT TCGTGAATGA  1260
CACGAAGGCC CCAGACTGGC AGGGCTACTT CTACACCGTG CTGCTGTTTG TCACTGCCTG  1320
CCTGCAGACC CTCGTGCTGC ACCAGTACTT CCACATCTGC TTCGTCAGTG GCATGAGGAT  1380
CAAGACCGCT GTCATTGGGG CTGTCTATCG GAAGGCCCTG GTGATCACCA ATTCAGCCAG  1440
AAAATCCTCC ACGGTCGGGG AGATTGTCAA CCTCATGTCT GTGGACGCTC AGAGGTTCAT  1500
GGACTTGGCC ACGTACATTA AGATGATCTG GTCAGCCCCC CTGCAAGTCA TCCTTGCTCT  1560
CTACCTCCTG TGGCTGAATC TGGGCCCTTC CGTCCTGGCT GGAGTGGCGG TGATGGTCCT  1620
CATGGTGCCC GTCAATGCTG TGATGGCGAT GAAGACCAAG ACGTATCAGG TGGCCCACAT  1680
GAAGAGCAAA GACAATCGGA TCAAGCTGAT GAACGAAATT CTCAATGGGA TCAAAGTGCT  1740
AAAGCTTTAT GCCTGGGAGC TGGCATTCAA GGACAAGGTG CTGGCCATCA GGCAGGAGGA  1800
GCTGAAGGTG CTGAAGAAGT CTGCCTACCT GTCAGCCGTG GGCACCTTCA CCTGGGTCTG  1860
CACGCCCTTT CTGGTGGCCT TGTGCACATT TGCCGTCTAC GTGACCATTG ACGAGAACAA  1920
CATCCTGGAT GCCCAGACAG CCTTCGTGTC TTTGGCCTTG TTCAACATCC TCCGGTTTCC  1980
CCTGAACATT CTCCCCATGG TCATCAGCAG CATCGTGCAG GCGAGTGTCT CCCTCAAACG  2040
CCTGAGGATC TTTCTCTCCC ATGAGGAGCT GGAACCTGAC AGCATCGAGC GACGGCCTGT  2100
CAAAGACGGC GGGGGCACGA ACAGCATCAC CGTGAGGAAT GCCACATTCA CCTGGGCCAG  2160
GAGCGACCCT CCCACACTGA ATGGCATCAC CTTCTCCATC CCCGAAGGTG CTTTGGTGGC  2220
```

```
CGTGGTGGGC CAGGTGGGCT GCGGAAAGTC GTCCCTGCTC TCAGCCCTCT TGGCTGAGAT  2280
GGACAAAGTG GAGGGGCACG TGGCTATCAA GGGCTCCGTG GCCTATGTGC CACAGCAGGC  2340
CTGGATTCAG AATGATTCTC TCCGAGAAAA CATCCTTTTT GGATGTCAGC TGGAGGAACC  2400
ATATTACAGG TCCGTGATAC AGGCCTGTGC CCTCCTCCCA GACCTGGAAA TCCTGCCCAG  2460
TGGGGATCGG ACAGAGATTG GCGAGAAGGG CGTGAACCTG TCTGGGGGCC AGAAGCAGCG  2520
CGTGAGCCTG GCCCGGCCG TGTACTCCAA CGCTGACATT TACCTCTTCG ATGATCCCCT  2580
CTCAGCAGTG GATGCCCATG TGGGAAAACA CATCTTTGAA AATGTGATTG GCCCCAAGGG  2640
GATGCTGAAG AACAAGACGC GGATCTTGGT CACGCACAGC ATGAGCTACT TGCCGCAGGT  2700
GGACGTCATC ATCGTCATGA GTGGCGGCAA GATCTCTGAG ATGGGCTCCT ACCAGGAGCT  2760
GCTGGCTCGA GACGGCGCCT TCGCTGAGTT CCTGCGTACC TATGCCAGCA CAGAGCAGGA  2820
GCAGGATGCA GAGGAGAACG GGGTCACGGG CGTCAGCGGT CCAGGGAAGG AAGCAAAGCA  2880
AATGGAGAAT GGCATGCTGG TGACGGACAG TGCAGGGAAG CAACTGCAGA GACAGCTCAG  2940
CAGCTCCTCC TCCTATAGTG GGGACATCAG CAGGCACCAC AACAGCACCG CAGAACTGCA  3000
GAAAGCTGAG GCCAAGAAGG AGGAGACCTG GAAGCTGATG GAGGCTGACA AGGCGCAGAC  3060
AGGGCAGGTC AAGCTTTCCG TGTACTGGGA CTACATGAAG GCCATCGGAC TCTTCATCTC  3120
CTTCCTCAGC ATCTTCCTTT TCATGTGTAA CCATGTGTCT GCGCTGGCTT CCAACTATTG  3180
GCTCAGCCTC TGGACTGATG ACCCCATCGT CAACGGGACT CAGGAGCACA CGAAAGTCCG  3240
GCTGAGCGTC TATGGAGCCC TGGGCATTTC ACAAGGGATC GCCGTGTTTG GCTACTCCAT  3300
GGCCGTGTCC ATCGGGGGGA TCTTGGCTTC CCGCTGTCTG CACGTGGACC TGCTGCACAG  3360
CATCCTGCGG TCACCCATGA GCTTCTTTGA GCGGACCCCC AGTGGGAACC TGGTGAACCG  3420
CTTCTCCAAG GAGCTGGACA CAGTGGACTC CATGATCCCG GAGGTCATCA AGATGTTCAT  3480
GGGCTCCCTG TTCAACGTCA TTGGTGCCTG CATCGTTATC CTGCTGGCCA CGCCCATCGC  3540
CGCCATCATC ATCCCGCCCC TTGGCCTCAT CTACTTCTTC GTCCAGAGGT TCTACGTGGC  3600
TTCCTCCCGG CAGCTGAAGC GCCTCGAGTC GGTCAGCCGC TCCCCGGTCT ATTCCCATTT  3660
CAACGAGACC TTGCTGGGGG TCAGCGTCAT TCGAGCCTTC GAGGAGCAGG AGCGCTTCAT  3720
CCACCAGAGT GACCTGAAGG TGGACGAGAA CCAGAAGGCC TATTACCCCA GCATCGTGGC  3780
CAACAGGTGG CTGGCCGTGC GGCTGGAGTG TGTGGGCAAC TGCATCGTTC TGTTTGCTGC  3840
CCTGTTTGCG GTGATCTCCA GGCACAGCCT CAGTGCTGGC TTGGTGGGCC TCTCAGTGTC  3900
TTACTCATTG CAGGTCACCA CGTACTTGAA CTGGCTGGTT CGGATGTCAT CTGAAATGGA  3960
AACCAACATC GTGGCCGTGG AGAGGCTCAA GGAGTATTCA GAGACTGAGA AGGAGGCGCC  4020
CTGGCAAATC CAGGAGACAG CTCCGCCCAG CAGCTGGCCC CAGGTGGGCC GAGTGGAATT  4080
CCGGAACTAC TGCCTGCGCT ACCGAGAGGA CCTGGACTTC GTTCTCAGGC ACATCAATGT  4140
CACGATCAAT GGGGGAGAAA AGGTCGGCAT CGTGGGGCGG ACGGGGAGCTG GGAAGTCGTC  4200
CCTGACCCTG GGCTTATTTC GGATCAACGA GTCTGCCGAA GGAGAGATCA TCATCGATGG  4260
CATCAACATC GCCAAGATCG GCCTGCACGA CCTCCGCTTC AAGATCACCA TCATCCCCCA  4320
GGACCCTGTT TTGTTTTCGG GTTCCCTCCG AATGAACCTG GACCCATTCA GCCAGTACTC  4380
GGATGAAGAA GTCTGGACGT CCCTGGAGCT GGCCCACCTG AAGGACTTCG TGTCAGCCCT  4440
TCCTGACAAG CTAGACCATG AATGTGCAGA AGGCGGGGAG AACCTCAGTG TCGGGCAGCG  4500
CCAGCTTGTG TGCCTAGCCC GGGCCCTGCT GAGGAAGACG AAGATCCTTG TGTTGGATGA  4560
GGCCACGGCA GCCGTGGACC TGGAAACGGA CGACCTCATC CAGTCCACCA TCCGGACACA  4620
GTTCGAGGAC TGCACCGTCC TCACCATCGC CCACCGGCTC AACACCATCA TGGACTACAC  4680
AAGGGTGATC GTCTTGGACA AAGGAGAAAT CCAGGAGTAC GGCGCCCCAT CGGACCTCCT  4740
GCAGCAGAGA GGTCTTTTCT ACAGCATGGC CAAAGACGCC GGCTTGGTGT GAGCCCCAGA  4800
GCTGGCATAT CTGGTCAGAA CTGCAGGGCC TATATGCCGA CGCCCAGGGA GGAGTCAGTA  4860
CCCCTGGTAA ACCAAGCCTC CCACACTGAA ACCAAAACAT AAAAACCAAA CCCAGACAAC  4920
CAAAACATAT TCAAAGCAGC AGCCACCGCC ATCCGGTCCC CTGCCTGGAA CTGGCTGTGA  4980
AGACCCAGGA GAGACAGAGA TGCGAACCAC C
```

SEQ ID NO:22 OBH2 Protein sequence:
Protein Accession #:   AAB46616

```
1          11         21         31         41         51
|          |          |          |          |          |
MALRGFCSAD GSDPLWDWNV TWNTSNPDFT KCFQNTVLVW VPCFYLWACF PFYFLYLSRH   60
DRGYIQMTPL NKTKTALGFL LWIVCWADLF YSFWERSRGI FLAPVFLVSP TLLGITTLLA  120
TFLIQLERRK GVQSSGIMLT FWLVALVCAL AILRSKIMTA LKEDAQVDLF RDITFYVYFS  180
LLLIQLVLSC FSDRSPLFSE TIHDPNPCPE SSASFLSRIT FWWITGLIVR GYRQPLEGSD  240
LWSLNKEDTS EQVVPVLVKN WKKECAKTRK QPVKVVYSSK DPAQPKESSK VDANEEVEAL  300
IVKSPQKEWN PSLFKVLYKT FGPYFLMSFF FKAIHDLMMF SGPQILKLLI KFVNDTKAPD  360
WQGYFYTVLL FVTACLQTLV LHQYFHICFV SGMRIKTAVI GAVYRKALVI TNSARKSSTV  420
GEIVNLMSVD AQRFMDLATY INMIWSAPLQ VILALYLLWL NLGPSVLAGV AVMVLMVPVN  480
AVMAMKTKTY QVAHMKSKDN RIKLMNEILN GIKVLKLYAW ELAFKDKVLA IRQEELKVLK  540
KSAYLSAVGT FTWVCTPFLV ALCTFAVYVT IDENNILDAQ TAFVSLALFN ILRFPLNILP  600
MVISSIVQAS VSLKRLRIFL SHEELEPDSI ERRPVKDGGG TNSITVRNAT FTWARSDPPT  660
LNGITFSIPE GALVAVVGQV GCGKSSLLSA LLAEMDKVEG HVAIKGSVAY VPQQAWIQND  720
SLRENILFGC QLEEPYYRSV IQACALLPDL EILPSGDRTE IGEKGVNLSG GQKQRVSLAR  780
AVYSNADIYL FDDPLSAVDA HVGKHIFENV IGPKGMLKNK TRILVTHSMS YLPQVDVIIV  840
MSGGKISEMG SYQELLARDG AFAEFLRTYA STEQEQDAEE NGVTGVSGPG KEAKQMENGM  900
LVTDSAGKQL QRQLSSSSSY SGDISRHHNS TAELQKAEAK KEETWKLMEA DKAQTGQVKL  960
SVYWDYMKAI GLFISFLSIF LFMCNHVSAL ASNYWLSLWT DDPIVNGTQE HTKVRLSVYG 1020
ALGISQGIAV FGYSMAVSIG GILASRCLHV DLLHSILRSP MSFFERTPSG NLVNRFSKEL 1080
DTVDSMIPEV IKMFMGSLFN VIGACIVILL ATPIAAIIIP PLGLIYFFVQ RFYVASSRQL 1140
KRLESVSRSP VYSHFNETLL GVSVIRAFEE QERFIHQSDL KVDENQKAYY PSIVANRWLA 1200
VRLECVGNCI VLFAALFAVI SRHSLSAGLV GLSVSYSLQV TTYLNWLVRM SSEMETNIVA 1260
VERLKEYSET EKEAPWQIQE TAPPSSWPQV GRVEFRNYCL RYREDLDFVL RHINVTINGG 1320
EKVGIVGRTG AGKSSLTLGL FRINESAEGE IIIDGINIAK IGLHDLRFKI TIIPQDPVLF 1380
SGSLRMNLDP FSQYSDEEVW TSLELAHLKD FVSALPDKLD HECAEGGENL SVGQRQLVCL 1440
ARALLRKTKI LVLDEATAAV DLETDDLIQS TIRTQFEDCT VLTIAHRLNT IMDYTRVIVL 1500
DKGEIQEYGA PSDLLQQRGL FYSMAKDAGL V
```

309

SEQ ID NO:23 PAA2 DNA SEQUENCE

Nucleic Acid Accession #: NM_013309
Coding sequence: 1-1290 (underlined sequences correspond to start and stop codons)

```
     1         11        21        31        41        51
     |         |         |         |         |         |
ATGGCCGGCT CTGGCGCGTG GAAGCGCCTC AAATCTATGC TAAGGAAGGA TGATGCGCCG    60
CTGTTTTTAA ATGACACCAG CGCCTTTGAC TTCTCGGATG AGGCGGGGGA CGAGGGGCTT   120
TCTCGGTTCA ACAAACTTCG AGTTGTGGTG GCCGATGACG GTTCCGAAGC CCCGGAAAGG   180
CCTGTTAACG GGGCGCACCC GACCCTCCAG GCCGACGATG ATTCCTTACT GGACCAAGAC   240
TTACCTTTGA CCAACAGTCA GCTGAGTTTG AAGGTGGACT CCTGTGACAA CTGCAGCAAA   300
CAGAGAGAGA TACTGAAGCA GAGAAAGGTG AAAGCCAGGT TGACCATTGC TGCCGTTCTG   360
TACTTGCTTT TCATGATTGG AGAACTTGTA GGTGGATACA TTGCAAATAG CCTAGCAATC   420
ATGACAGATG CACTTCATAT GTTAACTGAC CTAAGCGCCA TCATACTCAC CCTGCTTGCT   480
TTGTGGCTAT CATCAAAATC ACCAACCAAA AGATTCACCT TTGGATTTCA TCGCTTAGAG   540
GTTTTGTCAG CTATGATTAG TGTGCTGTTG GTGTATATAC TTATGGGATT CCTCTTATAT   600
GAAGCTGTGC AAAGAACTAT CCATATGAAC TATGAAATAA ATGGAGATAT AATGCTCATC   660
ACCGCAGCTG TTGGAGTTGC AGTTAATGTA ATAATGGGGT TTCTGTTGAA CCAGTCTGGT   720
CACCGTCACT CCCATTCCCA CTCCCTGCCT TCAAATTCCC CTACCAGAGG TTCTGGGTGT   780
GAACGTAACC ATGGGCAGGA TAGCCTGGCA GTGAGAGCTG CATTTGTACA TGCTTTGGGA   840
GATTTGGTAC AGAGTGTTGG TGTGCTAATA GCTGCATACA TCATACGATT CAAGCCAGAA   900
TACAAGATTG CTGATCCCAT CTGTACATAC GTATTTTCAT TACTTGTGGC TTTTACAACA   960
TTTCGAATCA TATGGGATAC AGTAGTTATA ATACTAGAAG GTGTGCCAAG CCATTTGAAT  1020
GTAGACTATA TCAAAGAAGC CTTGATGAAA ATAGAAGATG TATATTCAGT CGAAGATTTA  1080
AATATCTGGT CTCTCACTTC AGGAAAATCT ACTGCCATAG TTCACATACA GCTAATTCCT  1140
GGAAGTTCAT CTAAATGGGA GGAAGTACAG TCCAAAGCAA ACCATTTATT ATTGAACACA  1200
TTTGGCATGT ATAGATGTAC TATTCAGCTT CAGAGTTACA GGCAAGAAGT GGACAGAACT  1260
TGTGCAAATT GTCAGAGTTC TAGTCCCTGA                                   
```

SEQ ID NO:24 PAA2 Protein sequence:
Protein Accession #: NP_037441

```
     1         11        21        31        41        51
     |         |         |         |         |         |
MAGSGAWKRL KSMLRKDDAP LFLNDTSAFD FSDEAGDEGL SRFNKLRVVV ADDGSEAPER    60
PVNGAHPTLQ ADDDSLLDQD LPLTNSQLSL KVDSCDNCSK QREILKQRKV KARLTIAAVL   120
YLLFMIGELV GGYIANSLAI MTDALHMLTD LSAIILTLLA LWLSSKSPTK RFTFGFHRLE   180
VLSAMISVLL VYILMGFLLY EAVQRTIHMN YEINGDIMLI TAAVGVAVNV IMGFLLNQSG   240
HRHSHSHSLP SNSPTRGSGC ERNHGQDSLA VRAAFVHALG DLVQSVGVLI AAYIIRFKPE   300
YKIADPICTY VFSLLVAFTT FRIIWDTVVI ILEGVPSHLN VDYIKEALMK IEDVYSVEDL   360
NIWSLTSGKS TAIVHIQLIP GSSSKWEEVQ SKANHLLLNT FGMYRCTIQL QSYRQEVDRT   420
CANCQSSSP                                                           
```

SEQ ID NO:25 PAA3 DNA SEQUENCE

Nucleic Acid Accession #: AB037765
Coding sequence: 375-2798 (underlined sequences correspond to start and stop codons)

```
     1         11        21        31        41        51
     |         |         |         |         |         |
GCCGAGTCGG TGGCGGCTGC AGGCTGGGAG GGAGAAGTGC TACGCCTTTG CAGGTTGGCG    60
AAGTGGTTCC AGGCTACCCG GCTAGTCTGG CACGGCCCCG TCTTCTGCCT CCTCCTCCGT   120
CGCGTGGCGG CGGGAACTGT TGGCCGCGCG GCCTCGGGAA CGGCCCAGGT CCCCGCCCGC   180
AGGTCCCGGG CAGATAACAT AGATCATCAG TAGAAAACTT CTTGAAGTTG TTCAAGAAAA   240
ATTTGAAAGT AGCAAAATAG AAAATAAAGA ATTAACAGCA GATACAGAGG ACAGCATGGA   300
AGTGTGTGTCT TAGGAAACAG AACACAGCAG TGAAAAAACA GACAAAATCC GCTCAGATAC   360
AACTGCAGCT GATAATGTTT TCCGGCTTCA ATGTCTTTAG AGTTGGGATC TCTTTTGTCA   420
TAATGTGCAT TTTTTACATG CCAACAGTAA ACTCTTTACC AGAACTGAGT CCTCAGAAAT   480
ATTTTAGTAC ATTGCAACCA GGTCTTGAAG AACTGAATGA GGCTGTTAGA CCTCTGCAGG   540
ACTATGGAAT TTCAGTTGCC AAGGTTAATT GTGTCAAAGA AGAAATATCA AGATACTGTG   600
GAAAAGAAAA GGATTTGATG AAAGCATATT TATTCAAGGG CAACATATTG CTCAGAGAAT   660
TCCCTACTGA CACCTTGTTT GATGTGAATG CCATTGTCGC CCATGTTCTC TTTGCTCTTC   720
TTTTTAGTGA AGTGAAATAT ATTACCAACC TGGAAGACCT TCAGAACATA GAAAATGCTC   780
TGAAAGGAAA AGCAAATATT ATATTCTCAT ATGTAAGAGC CATTGGAATA CCAGAGCACA   840
GAGCAGTCAT GGAAGCCGGT TTTGTGTATG GGACTACATA CCAATTTGTC TTAACCACAG   900
AAAATTGCCCT TTTGGAAAGT ATTGGCTCTG AGGATGTGGA ATATGCACAT CTCTACTTTT   960
TTCATTGTAA ACTAGTCTTG GACTTGACCC AGCAATGTAG AAGAACACTA ATGGAACAGC  1020
CATTGACTAC ACTGAACATT CACCTGTTTA TTAAGCAAT GAAAGCACCT CTGTTGACTG  1080
AAGTTGCTGA AGATCCTCAA CAAGTTTCAA CTGTCCATCT CCAACTGGGC TTACCACTGG  1140
TTTTTATTGT TAGCCAACAG GCTACTTATG AAGCTGATAG AGAACTGCA GAATGGGTTG  1200
CTTGGCGTCT TCTGGGAAAA GCAGGAGTTC TACTCTTGTT AAGGGACTCT TTGGAAGTGA  1260
ACATTCCTCA AGATGCTAAT GTGGTCTTCA AAAGAGCAGA AGAGGGAGTT CCAGTGGAAT  1320
TTTTGGTATT ACATGATGTT GATTTAATAA TATCTCATGT GGAAAATAAT ATGCACATTG  1380
AGGAAATACA AGAAGATGAA GACAATGACA TGGAAGGTCC AGATATAGAT GTTCAGGATG  1440
ATGAAGTGGC AGAAACTGTT TTCAGAGATA GGAAGAGAAA ATTACCTTTG GAACTTACAG  1500
```

```
TGGAACTAAC AGAAGAAACA TTTAATGCAA CAGTGATGGC TTCTGACAGC ATAGTACTCT    1560
TCTATGCTGG TTGGCAAGCA GTATCCATGG CATTTTTGCA ATCCTATATT GATGTGGCAG    1620
TTAAACTGAA AGGCACATCT ACTATGCTTC TTACTAGAAT AAACTGTGCA GATTGGTCTG    1680
ATGTATGTAC TAAGCAAAAT GTTACTGAAT TTCCTATCAT AAAGATGTAC AAGAAAGGCG    1740
AGAACCCAGT ATCTTATGCT GGAATGTTAG GAACCAAAGA TCTCCTAAAA TTTATCCAGC    1800
TCAACAGGAT TTCATATCCA GTGAATATAA CATCGATCCA AGAAGCAGAA GAATATTTAA    1860
GTGGGGAATT ATATAAAGAC CTCATCTTGT ATTCTAGTGT GTCAGTATTG GGACTATTTA    1920
GTCCAACCAT GAAAACAGCA AAAGAAGATT TTAGTGAAGC AGGAAACTAC CTAAAAGGAT    1980
ATGTTATCAC TGGAATTTAT TCTGAAGAAG ATGTTTTGCT ACTGTCAACC AAATATGCTG    2040
CAAGTCTTCC AGCCCTGCTG CTTGCCAGAC ACACAGAAGG CAAAATAGAG AGCATCCCAC    2100
TAGCTAGCAC ACATGCACAA GACATAGTTC AAATAATAAC AGATGCACTA CTGGAAATGT    2160
TTCCGGAAAT CACTGTGGAA AATCTTCCCA GTTATTTCAG ACTTCAGAAA CCATTATTGA    2220
TTTTGTTCAG TGATGGCACT GTAAATCCTC AATATAAAAA AGCAATATTG ACACTGGTAA    2280
AGCAGAAATA CTTGGATTCA TTTACTCCAT GCTGGTTAAA TCTAAAGAAT ACTCCAGTGG    2340
GGAGAGGAAT CTTGCGGGCA TATTTGATC CTCTGCCTCC CCTTCCTCTT CTTGTTTTGG     2400
TGAATCTGCA TTCAGGTGGC CAAGTATTTG CATTTCCTTC AGACCAGGCT ATAATTGAAG    2460
AAAACCTTGT ATTGTGGCTG AAGAAATTAG AAGCAGGACT AGAAAATCAT ATCACAATTT    2520
TACCTGCTCA AGAATGGAAA CCTCCTCTTC CAGCTTATGA TTTTCTAAGT ATGATAGATG    2580
CCGCAACATC TCAACGTGGC ACTAGGAAAG TTCCCAAGTG TATGAAAGAA ACAGATGTGC    2640
AGGAGAATGA TAAGGAACAA CATGAAGATA AATCGGCAGT CAGAAAAGAA CCGATTGAAA    2700
CTCTGAGAAT AAAGCATTGG AATAGAAGTA ATTGGTTTAA AGAAGCAGAA AAATCATTTA    2760
GACGTGATAA AGAGTTAGGA TGCTCAAAAG TGAACTAATT TTATAGGGCT GTGGTTTCCA    2820
AAATTTTTTT GGCATGATAG ACTTAATTTA TTTCCTTAAA GAATAATATT AAATCATTTC    2880
AAGTTTGCAG ACTAGTGCCA TCCAATAGAA TTATAATATA AGTCACATAT TTTATTTAAA    2940
ATTTTCTAGT AACTACATTA AACAAAGTAA AAGTGAGCAG GGCAAAATAA TTTTGATATT    3000
ACTTTTCACC CAGTAGTATA CCCAAAATAG CGAAATATAG AAATTATTAA TGAGATATTT    3060
TACATCCTTT TTTGTACCAA GTCTTCTAAA TGCAGTACAT ATTTTATACT TACTGCATTT    3120
CTTACTTCCG AGTAGCCATA TTTCAAGTGT TCATTGCCAC ATGTGGCCTG TGACTACTGT    3180
ATTGGACAGT TCAGTACTAG ACAAAAACTA GCATAATTAA CTTAGTTCTA GCCATGATTT    3240
CTATTTGGAT TAAAATTAAA CTCTAATCAC AGTTAACTCC ACAGTGCATT CATGCAGCTG    3300
ACAGTTATAT TTGTTTTATT GGAGTCATGA TATTAAAATC AGCGTTTGTC AACCTCAGGG    3360
GATATTTAGC AATTGTCGGG AGACATTTTT GATGTCATGA CTAGGGCAGT TATTGACATT    3420
TAGTGAGTAG AGGCCATGGA TCCTGCTAAA TAACCTGCAT TGGACAGCGC CCCACAACAA    3480
AGAATTATCC TGCCCGAAAT GGTAGTCGTG CCAAGGCTGA GTAACCTTGT GTTAAAAGTA    3540
ACCTGTGGCA GACTAGGTTT CCAGAATTTC CTGGTTCTGC TCACGTATCA TGTTTGAAAA    3600
AATTTTGGCT ATTAAAGATA TGTATTAGAT GGTCTTATCC TGATTATTAC CTGGATACAA    3660
CTTGATCTTT TCTAATATTT TCAGAAAGTG ATGGGATAAC CCTAGAAGAG GACTCAGAAT    3720
GATATTTATA TTTTAAGTGA GTCTTAAAAC CTCCTCTTAT TTCTACAAGT TATATGGCTA    3780
AATTTCAGAT TGAACAGGGA TTCAGCATTC TGCCATCTCC TCATGGAAAG AGAGGCTCCC    3840
TCATCTGAAG CGTCTCTGAA ATCTACCCTT GCAAGCTTCA GACAAATCAG TTGATCTCCC    3900
TGAGCCACAC GGCCTCATTC TGTGAGGGAG GGAAAGATTA GCCAAAGAGT TAATTTTCAT    3960
TCCAAATCAC TTAGCTGTTA GACTGATCTG TTTGTAGCAG TTGTTTGTCT CATTTTTGCT    4020
CTGTGCATTT TTTGAGACAT TTGTTGAGAA TATTCTATTT GGTGCTCTAC TGTATTTTTC    4080
TTTTTAATAT CTACTTGATA TCTTGTTCTT TAAATTTTCT TCACATATGG TTTGCCTGAT    4140
ACAACTGATT TTTATAACTG AAATTTAAGG AATCTAACAG CTAAAACTCA GTAAGTGCAT    4200
MTATTTCCTT ATAACATAGA CCCGTTGCTA CTCTCAGCAC CCTCTCCTCA ATTTTTTTTC    4260
CTGTAGCATG TGATGCCTGA TTAAACTCAT TTTCATTTGC TTTTATTTCT AATATGGGAA    4320
CAATGAGAGT GAACTCTAAA TATAGGTTGT AGTAATAAAA CATCATTAGC CTAATTATTA    4380
GAAAATGCTA ATTAAGTACC AGCACATAGA AACATGAAAT TGCTTAGTCA TTGTACCTTT    4440
GTCAGCAATT TTGACAGTCA TTAATGTTTG TCATAATTTT AAATAAAGTG TCTGGGTTTC    4500
AGAATACCTT CAAAAAAAAA AAAAAA
```

SEQ ID NO:26 PAA3 Protein sequence:
Protein Accession #:   BAA92582

```
1          11         21         31         41         51
|          |          |          |          |          |
MFSGFNVFRV GISFVIMCIF YMPTVNSLPE LSPQKYFSTL QPGLEELNEA VRPLQDYGIS    60
VAKVNCVKEE ISRYCGKEKD LMKAYLFKGN ILLREFPTDT LFDVNAIVAH VLFALLFSEV   120
KYITNLEDLQ NIENALKGKA NIIFSYVRAI GIPEHRAVME AGFVYGTTYQ FVLTTEIALL   180
ESIGSEDVEY AHLYFFHCKL VLDLTQQCRR TLMEQPLTTL NIHLFIKTMK APLLTEVAED   240
PQQVSTVHLQ LGLPLVFIVS QQATYEADRR TAEWVAWRLL GKAGVLLLLR DSLEVNIPQD   300
ANVVFKRAEE GVPVEFLVLH DVDLIISHVE NNMHIEEIQE DEDNDMEGPD IDVQDDEVAE   360
TVFRDRKRKL PLELTVELTE ETFNATVMAS DSIVLFYAGW QAVSMAFLQS YIDVAVKLKG   420
TSTMLLTRIN CADWSDVCTK QNVTEFPIIK MYKKGENPVS YAGMLGTKDL LKFIQLNRIS   480
YPVNITSIQE AEEYLSGELY KDLILYSSVS VLGLFSPTMK TAKEDFSEAG NYLKGYVITG   540
IYSEEDVLLL STKYAASLPA LLLARHTEGK IESIPLASTH AQDIVQIITD ALLEMFPEIT   600
VENLPSYFRL QKPLLILFSD GTVNPQYKKA ILTLVKQKYL DSFTPCWLNL KNTPVGRGIL   660
RAYFDPLPPL PLLVLVNLHS GGQVFAFPSD QAIIEENLVL WLKKLEAGLE NHITILPAQE   720
WKPPLPAYDF LSMIDAATSQ RGTRKVPKCM KETDVQENDK EQHEDKSAVR KEPIETLRIK   780
HWNRSNWFKE AEKSFRRDKE LGCSKVN
```

SEQ ID NO:27 PAA5 DNA SEQUENCE

Nucleic Acid Accession #:   NM_012449
Coding sequence: 66-1085 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
CCGAGACTCA CGGTCAAGCT AAGGCGAAGA GTGGGTGGCT GAAGCCATAC TATTTTATAG     60
AATTAATGGA AAGCAGAAAA GACATCACAA ACCAAGAAGA ACTTTGGAAA ATGAAGCCTA    120
```

311

```
GGAGAAATTT AGAAGAAGAC GATTATTTGC ATAAGGACAC GGGAGAGACC AGCATGCTAA  180
AAAGACCTGT GCTTTTGCAT TTGCACCAAA CAGCCCATGC TGATGAATTT GACTGCCCTT  240
CAGAACTTCA GCACACACAG GAACTCTTTC CACAGTGGCA CTTGCCAATT AAAATAGCTG  300
CTATTATAGC ATCTCTGACT TTTCTTTACA CTCTTCTGAG GGAAGTAATT CACCCTTTAG  360
CAACTTCCCA TCAACAATAT TTTTATAAAA TTCCAATCCT GGTCATCAAC AAAGTCTTGC  420
CAATGGTTTC CATCACTCTC TTGGCATTGG TTTACCTGCC AGGTGTGATA GCAGCAATTG  480
TCCAACTTCA TAATGGAACC AAGTATAAGA AGTTCCACA TTGGTTGGAT AAGTGGATGT  540
TAACAAGAAA GCAGTTTGGG CTTCTCAGTT TCTTTTTTGC TGTACTGCAT GCAATTTATA  600
GTCTGTCTTA CCCAATGAGG CGATCCTACA GATACAAGTT GCTAAACTGG GCATATCAAC  660
AGGTCCAACA AAATAAAGAA GATGCCTGGA TTGAGCATGA TGTTTGGAGA ATGGAGATTT  720
ATGTGTCTCT GGGAATTGTG GGATTGGCAA TACTGGCTCT GTTGGCTGTG ACATCTATTC  780
CATCTGTGAG TGACTCTTTG ACATGGAGAG AATTTCACTA TATTCAGAGC AAGCTAGGAA  840
TTGTTTCCCT TCTACTGGGC ACAATACACG CATTGATTTT TGCCTGGAAT AAGTGGATAG  900
ATATAAAACA ATTTGTATGG TATACACCTC CAACTTTTAT GATAGCTGTT TTCCTTCCAA  960
TTGTTGTCCT GATATTTAAA AGCATACTAT TCCTGCCATG CTTGAGGAAG AAGATACTGA  1020
AGATTAGACA TGGTTGGGAA GACGTCACCA AAATTAACAA AACTGAGATA TGTTCCCAGT  1080
TGTAGAATTA CTGTTTACAC ACATTTTTGT TCAATATTGA TATATTTTAT CACCAACATT  1140
TCAAGTTTGT ATTTGTTAAT AAAATGATTA TTCAAGGAAA AAAAAAAAAA AAAAA
```

SEQ ID NO:28 PAA5 Protein sequence
Protein Accession #:   NP_036581

```
1          11         21         31         41         51
|          |          |          |          |          |
MESRKDITNQ EELWKMKPRR NLEEDDYLHK DTGETSMLKR PVLLHLHQTA HADEFDCPSE   60
LQHTQELFPQ WHLPIKIAAI IASLTFLYTL LREVIHPLAT SHQQYFYKIP ILVINKVLPM  120
VSITLLALVY LPGVIAAIVQ LHNGTKYKKF PHWLDKWMLT RKQFGLLSFF FAVLHAIYSL  180
SYPMRRSYRY KLLNWAYQQV QQNKEDAWIE HDVWRMEIYV SLGIVGLAIL ALLAVTSIPS  240
VSDSLTWREF HYIQSKLGIV SLLLGTIHAL IFAWNKWIDI KQFVWYTPPT FMIAVFLPIV  300
VLIFKSILFL PCLRKKILKI RHGWEDVTKI NKTEICSQL
```

SEQ ID NO:29 PAA7 DNA SEQUENCE

Nucleic Acid Accession #:   NM_030774
Coding sequence: 1-963 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGAGTTCCT GCAACTTCAC ACATGCCACC TTTGTGCTTA TTGGTATCCC AGGATTAGAG   60
AAAGCCCATT TCTGGGTTGG CTTCCCCCTC CTTTCCATGT ATGTAGTGGC AATGTTTGGA  120
AACTGCATCG TGGTCTTCAT CGTAAGGACG GAACGCAGCC TGCACGCTCC GATGTACCTC  180
TTTCTCTGCA TGCTTGCAGC CATTGACCTG GCCTTATCCA CATCCACCAT GCCTAAGATC  240
CTTGCCCTTT TCTGGTTTGA TTCCCGAGAG ATTAGCTTTG AGGCCTGTCT TACCCAGATG  300
TTCTTTATTC ATGCCCTCTC AGCCATTGAA TCCACCATCC TGCTGGCCAT GGCCTTTGAC  360
CGTTATGTGG CCATCTGCCA CCCACTGCGC CATGCTGCAG TGCTCAACAA TACAGTAACA  420
GCCCAGATTG GCATCGTGGC TGTGGTCCGC GGATCCCTCT TTTTTTTCCC ACTGCCTCTG  480
CTGATCAAGC GGCTGGCCTT CTGCCACTCC AATGTCCTCT CGCACTCCTA TTGTGTCCAC  540
CAGGATGTAA TGAAGTTGGC CTATGCAGAC ACTTTGCCCA ATGTGGTATA TGGTCTTACT  600
GCCATTCTGC TGGTCATGGG CGTGGACGTA ATGTTCATCT CCTTGTCCTA TTTTCTGATA  660
ATACGAACGG TTCTGCAACT GCCTTCCAAG TCAGAGCGGG CCAAGGCCTT TGGAACCTGT  720
GTGTCACACA TTGGTGTGGT ACTCGCCTTC TATGTGCCAC TTATTGGCCT CTCAGTGGTA  780
CACCGCTTTG GAAACAGCCT TCATCCCATT GTGCGTGTTG TCATGGGTGA CATCTACCTG  840
CTGCTGCCTC CTGTCATCAA TCCCATCATC TATGGTGCCA AAACCAAACA GATCAGAACA  900
CGGGTGCTGG CTATGTTCAA GATCAGCTGT GACAAGGACT TGCAGGCTGT GGGGAGGCAAG  960
TGACCCTTAA CACTACACTT CTCCTTATCT TTATTGGCTT GATAAACATA ATTATTTCTA  1020
ACACTAGCTT ATTTCCAGTT GCCCATAAGC ACATCAGTAC TTTTCTCTGG CTGGAATAGT  1080
AAACTAAAGT ATGGTACATC TACCTAAAGG ACTATTATGT GGAATAATAC ATACTAATGA  1140
AGTATTACAT GATTTAAAGA CTACAATAAA ACCAAACATG CTTATAACAT TAAGAAAAAC  1200
AATAAAGATA CATGATTGAA ACCAAGTTGA AAAATAGCAT ATGCCTTGGA GGAAATGTGC  1260
TCAAATTACT AATGATTTAG TGTTGTCCCT ACTTTCTCTC TCTTTTTTCT TTCTTTTTTT  1320
TTTATTATGG TTAGCTGTCA CATACAACTT TTTTTTTTTT TGAGATGGGG TCTCGCTCTG  1380
TCACCAGGCT GGAGTGCAGT GGCGCGATCT CGGCTCACTG CAACCTCCAC ATCCCATGTT  1440
GAAGTAATTC TTCTGCCTCA GCCTCCCGAG TAGCTGGGAC TAGAGGAACG TGCCACCATG  1500
ACTGGCTAAT TTTCTGTATT TTTTAGTAGA GACAGAGTTT CACCATGTTG GCCAGGATGG  1560
TCTCGATCTC CTGACCTTGT GATCCACCCG CCTCAGCCTC CCAAAGTGTT GGGATTACAG  1620
GTGTGAACCA CTGTGCCCGG CCTGTGTACA ACTTTTTAAA TAGGGAATAT GATAGCTTCG  1680
CATGGTGGTG TGCACCTATA GCCCCCACTG CCTGGAAAGC TGAGGTGGGA GAATCGCTTG  1740
AGTCCAGGAG TTTGAGGTTA CAGTGATCCA CGATCGTACC ACTACACTCC AGCCTGGGCA  1800
ACAGAGCAAG ACCCTGTCTC AAAGCATAAA ATGGAATAAC ATATCAAATG AAACAGGGAA  1860
AATGAAGCTG ACAATTTATG GAAGCCAGGG CTTGTCACAG TCTCTACTGT TATTATGCAT  1920
TACCTGGGAA TTTATATAAG CCCTTAATAA TARTGCCAAT GAACATCTCA TGTGTGCTCA  1980
CAATGTTCTG GCACTATTAT AAGTGCTTCA CAGGTTTTAT GTGTTCTTCG TAACTTTATG  2040
GAGTAGGTAC CATTTGTGTC TCTTTATTAT AAGTGAGAGA AATGAAGTTT ATATTATCAA  2100
GGGGACTAAA GTCACACGGC TTGTGGGCAC TGTGCCAAGA TTTAAAATTA AATTTGATGG  2160
TTGAATACAG TTACTTAATG ACCATGTTAT ATTGCTTCCT GTGTAACATC TGCCATTTAT  2220
TTCCTCAGCT GTACAAATCC TCTGTTTTCT CTCTGTTACA CACTAACATC AATGGCTTTG  2280
TACTTGTGAT GAGAGATAAC CTTGCCCTAG TTGTGGGCAA CACATGCAGA ATAATCCTGT  2340
TTTACAGCTG CCTTTCGTGA TCTTATTGCT TGCTTTTTTC CAGATTCAGG GAGAATGTTG  2400
TTGTCTATTT GTCTCTTACA TCTCCTTGAT CATGTCTTCA TTTTTTAATG TGCTCTGTAC  2460
CTGTCAAAAA TTTTGAATGT ACACCACATG CTATTGTCTG AACTTGAGTA TAAGATAAAA  2520
TAAAATTTTA TTTTAAATTT T
```

312

SEQ ID NO:30 PAA7 PROTEIN SEQUENCE

Protein Accession #:   NP_110401

```
1          11         21         31         41         51
|          |          |          |          |          |
MSSCNFTHAT FVLIGIPGLE KAHFWVGFPL LSMYVVAMFG NCIVVFIVRT ERSLHAPMYL    60
FLCMLAAIDL ALSTSTMPKI LALFWFDSRE ISFEACLTQM FFIHALSAIE STILLAMAFD   120
RYVAICHPLR HAAVLNNTVT AQIGIVAVVR GSLFFFPLPL LIKRLAFCHS NVLSHSYCVH   180
QDVMKLAYAD TLPNVVYGLT AILLVMGVDV MFISLSYFLI IRTVLQLPSK SERAKAFGTC   240
VSHIGVVLAF YVPLIGLSVV HRFGNSLHPI VRVVMGDIYL LLPPVINPII YGAKTKQIRT   300
RVLAMFKISC DKDLQAVGGK
```

SEQ ID NO:31 PAV6 DNA SEQUENCE

Nucleic Acid Accession #:   XM_050837
Coding sequence: 1-1020 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGAACTGGG AGCTGCTGCT GTGGCTGCTG GTGCTGTGCG CGCTGCTCCT GCTCTTGGTG    60
CAGCTGCTGC GCTTCCTGAG GGCTGACGGC GACCTGACGC TACTATGGGC CGAGTGGCAG   120
GGACGACGCC CAGAATGGGA GCTGACTGAT ATGGTGGTGT GGGTGACTGG AGCCTCGAGT   180
GGAATTGGTG AGGAGCTGGC TTACCAGTTG TCTAAACTAG GAGTTTCTCT TGTGCTGTCA   240
GCCAGAAGAG TGCATGAGCT GGAAAGGGTG AAAAGAAGAT GCCTAGAGAA TGGCAATTTA   300
AAAGAAAAAG ATATACTTGT TTTGCCCCTT GACCTGACCG ACACTGGTTC CCATGAAGCG   360
GCTACCAAAG CTGTTCTCCA GGAGTTTGGT AGAATCGACA TTCTGGTCAA CAATGGTGGA   420
ATGTCCCAGC GTTCTCTGTG CATGGATACC AGCTTGGATG TCTACAGAAA GCTAATAGAG   480
CTTAACTACT TAGGGACGGT GTCCTTGACA AAATGTGTTC TGCCTCACAT GATCGAGAGG   540
AAGCAAGGAA AGATTGTTAC TGTGAATAGC ATCCTGGGTA TCATATCTGT ACCTCTTTCC   600
ATTGGATACT GTGCTAGCAA GCATGCTCTC CGGGGTTTTT TTAATGGCCT TCGAACAGAA   660
CTTGCCCACAT ACCCAGGTAT AATAGTTTCT AACATTTGCC CAGGACCTGT GCAATCAAAT   720
ATTGTGGAGA ATTCCCTAGC TGGAGAAGTC ACAAAGACTA TAGGCAATAA TGGAGACCAG   780
TCCCACAAGA TGACAACCAG TCGTTGTGTG CGGCTGATGT TAATCAGCAT GGCCAATGAT   840
TTGAAAGAAG TTTGGATCTC AGAACAACCT TTCTTGTTAG TAACATATTT GTGGCAATAC   900
ATGCCAACCT GGGCCTGGTG GATAACCAAC AAGATGGGGA AGAAAAGGAT TGAGAACTTT   960
AAGAGTGGTG TGGATGCAGA CTCTTCTTAT TTTAAAATCT TTAAGACAAA ACATGACTGA
```

SEQ ID NO:32 PAV6 Protein sequence

Protein Accession #:   XP_050837

```
1          11         21         31         41         51
|          |          |          |          |          |
MNWELLLWLL VLCALLLLLV QLLRFLRADG DLTLLWAEWQ GRRPEWELTD MVVWVTGASS    60
GIGEELAYQL SKLGVSLVLS ARRVHELERV KRRCLENGNL KEKDILVLPL DLTDTGSHEA   120
ATKAVLQEFG RIDILVNNGG MSQRSLCMDT SLDVYRKLIE LNYLGTVSLT KCVLPHMIER   180
KQGKIVTVNS ILGIISVPLS IGYCASKHAL RGFFNGLRTE LATYPGIIVS NICPGPVQSN   240
IVENSLAGEV TKTIGNNGDQ SHKMTTSRCV RLMLISMAND LKEVWISEQP FLLVTYLWQY   300
MPTWAWWITN KMGKKRIENF KSGVDADSSY FKIFKTKHD
```

SEQ ID NO:33 PBA6 DNA SEQUENCE

Nucleic Acid Accession #:   NM_006853
Coding sequence: 26-874 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
AGGAATCTGC GCTCGGGTTC CGCAGATGCA GAGGTTGAGG TGGCTGCGGG ACTGGAAGTC    60
ATCGGGCAGA GGTCTCACAG CAGCCAAGGA ACCTGGGGCC CGCTCCTCCC CCCTCCAGGC   120
CATGAGGATT CTGCAGTTAA TCCTGCTTGC TCTGGCAACA GGGCTTGTAG GGGGAGAGAC   180
CAGGATCATC AAGGGGTTCG AGTGCAAGCC TCACTCCCAG CCCTGGCAGG CAGCCCTGTT   240
CGAGAAGACG CGGCTACTCT GTGGGGCGAC GCTCATCGCC CCAGATGGC TCCTGACACG   300
AGCCCACTGC CTCAAGCCCC GCTACATAGT TCACCTGGGG CAGCACAACC TCCAGAAGGA   360
GGAGGGCTGT GAGCAGACCC GGACAGCCAC TGAGTCCTTC CCCCACCCCG GCTTCAACAA   420
CAGCCTCCCC AACAAAGACC ACCGCAATGA CATCATGCTG GTGAAGATGG CATCGCCAGT   480
CTCCATCACC TGGGCTGTGC GACCCCTCAC CCTCTCCTCA CGCTGTGTCA CTGCTGGCAC   540
CAGCTGCCTC ATTTCCGGCT GGGGCAGCAC GTCCAGCCCC CAGTTACGCC TGCCTCACAC   600
CTTGCGATGC GCCAACATCA CCATCATTGA GCACCAGAAG TGTGAGAACG CCTACCCCGG   660
CAACATCACA GACACCATGG TGTGTGCCAG CGTGCAGGAA GGGGGCAAGG ACTCCTGCCA   720
GGGTGACTCC GGGGGCCCTC TGGTCTGTAA CCAGTCTCTT CAAGGCATTA TCTCCTGGGG   780
CCAGGATCCG TGTGCGATCA CCCGAAAGCC TGGTGTCTAC ACGAAAGTCT GCAAATATGT   840
GGACTGGATC CAGGAGACGA TGAAGAACAA TTAGACTGGA CCCACCCACC ACAGCCCATC   900
ACCCTCCATT TCCACTTGGT GTTTGGTTCC TGTTCACTCT GTTAATAAGA AACCCTAAGC   960
CAAGACCCTC TACGAACATT CTTTGGGCCT CCTGGACTAC AGGAGATGCT GTCACTTAAT  1020
AATCAACCTG GGGTTCGAAA TCAGTGAGAC CTGGATTCAA ATTCTGCCTT GAAATATTGT  1080
GACTCTGGGA ATGACAACAC CTGGTTTGTT CTCTGTTGTA TCCCCAGCCC CAAAGACAGC  1140
TCCTGGCCAT ATATCAAGGT TTCAATAAAT ATTTGCTAAA TGAGTG
```

SEQ ID NO:34 PBA6 PROTEIN SEQUENCE

Protein Accession #:   NP_006844

313

```
        1          11         21         31         41         51
        |          |          |          |          |          |
5       MRILQLILLA LATGLVGGET RIIKGFECKP HSQPWQAALF EKTRLLCGAT LIAPRWLLTA     60
        AHCLKPRYIV HLGQHNLQKE EGCEQTRTAT ESFPHPGFNN SLPNKDHRND IMLVKMASPV    120
        SITWAVRPLT LSSRCVTAGT SCLISGWGST SSPQLRLPHT LRCANITIIE HQKCENAYPG    180
        NITDTMVCAS VQEGGKDSCQ GDSGGPLVCN QSLQGIISWG QDPCAITRKP GVYTKVCKYV    240
        DWIQETMKNN
```

SEQ ID NO:35 PBC1 DNA SEQUENCE

Nucleic Acid Accession #: NM_001775
Coding sequence: 70-972 (underlined sequences correspond to start and stop codons)

```
        1          11         21         31         41         51
        |          |          |          |          |          |
15      CTAAAGCTCT CTTGCTGCCT AGCCTCCTGC CGGCCTCATC TTCGCCCAGC CAACCCCGCC     60
        TGGAGCCCTA TGGCCAACTG CGAGTTCAGC CCGGTGTCCG GGGACAAACC CTGCTGCCGG    120
        CTCTCTAGGA GAGCCCAACT CTGTCTTGGC GTCAGTATCC TGGTCCTGAT CCTCGTCGTG    180
        GTGCTCGCGG TGGTCGTCCC GAGGTGGCGC CAGACGTGGA GCGGTCCGGG CACCACCAAG    240
        CGCTTTCCCG AGACCGTCCT GGCCGCGATGC GTCAAGTACA CTGAAATTCA TCCTGAGATG    300
        AGACATGTAG ACTGCCAAAG TGTATGGGAT GCTTTCAAGG GTGCATTTAT TTCAAAACAT    360
        CCTTGCAACA TTACTGAAGA AGACTATCAG CCACTAATGA AGTTGGGAAC TCAGACCGTA    420
        CCTTGCAACA AGATTCTTCT TTGGAGCAGA ATAAAAGATC TGGCCCATCA GTTCACACAG    480
        GTCCAGCGGG ACATGTTCAC CCTGGAGGAC ACGCTGCTAG GCTACCTTGC TGATGACCTC    540
        ACATGGTGTG GTGAATTCAA CACTTCCAAA ATAAACTATC AATCTTGCCC AGACTGGAGA    600
        AAGGACTGCA GCAACAACCC TGTTTCAGTA TTCTGGAAAA CGGTTTCCCG CAGGTTTGCA    660
        GAAGCTGCCT GTGATGTGGT CCATGTGATG CTCAATGGAT CCCGCAGTAA AATCTTTGAC    720
        AAAAACAGCA CTTTTGGGAG TGTGGAAGTC CATAATTTGC AACCAGAGAA GGTTCAGACA    780
        CTAGAGGCCT GGGTGATACA TGGTGGAAGA GAAGATTCCA GAGACTTATG CCAGGATCCC    840
        ACCATAAAAG AGCTGGAATC GATTATAAGC AAAAGGAATA TTCAATTTTC CTGCAAGAAT    900
        ATCTACAGAC CTGACAAGTT TCTTCAGTGT GTGAAAAATC CTGAGGATTC ATCTTGCACA    960
        TCTGAGATCT GAGCCAGTCG CTGTGGTTGT TTTAGCTCCT TGACTCCTTG TGGTTTATGT   1020
        CATCATACAT GACTCAGCAT ACCTGCTGGT GCAGAGCTGA AGATTTTGGA GGGTCCTCCA   1080
        CAATAAGGTC AATGCCAGAG ACGGAAGCCT TTTTCCCCAA AGTCTTAAAA TAACTTATAT   1140
        CATCAGCATA CCTTTATTGT GATCTATCAA TAGTCAAGAA AAATTATTGT ATAAGATTAG   1200
        AATGAAAATT GTATGTTAAG TTACTTCCTT TAG
```

SEQ ID NO:36 PBC1 Protein sequence
Protein Accession #: NP_001766

```
        1          11         21         31         41         51
        |          |          |          |          |          |
        MANCEFSPVS GDKPCCRLSR RAQLCLGVSI LVLILVVVLA VVVPRWRQTW SGPGTTKRFP     60
        ETVLARCVKY TEIHPEMRHV DCQSVWDAFK GAFISKHPCN ITEEDYQPLM KLGTQTVPCN    120
        KILLWSRIKD LAHQFTQVQR DMFTLEDTLL GYLADDLTWC GEFNTSKINY QSCPDWRKDC    180
        SNNPVSVFWK TVSRRFAEAA CDVVHVMLNG SRSKIFDKNS TFGSVEVHNL QPEKVQTLEA    240
        WVIHGGREDS RDLCQDPTIK ELESIISKRN IQFSCKNIYR PDKFLQCVKN PEDSSCTSEI
```

SEQ ID NO:37 PBH1 DNA SEQUENCE

Nucleic Acid Accession #: XM_017718
Coding sequence: 1-3315 (underlined sequences correspond to start and stop codons)

```
        1          11         21         31         41         51
        |          |          |          |          |          |
55      ATGTCCTTTC GGGCAGCCAG GCTCAGCATG AGGAACAGAA GGAATGACAC TCTGGACAGC     60
        ACCCGGACCC TGTACTCCAG CGCGTCTCGG AGCACAGACT TGTCTTACAG TGAAAGCGAC    120
        TTGGTGAATT TTATTCAAGC AAATTTTAAG AAACGAGAAT GTGTCTTCTT TACCAAAGAT    180
        TCCAAGGCCA CGGAGAATGT GTGCAAGTGT GGCTATGCCC AGAGCCAGCA CATGGAAGGC    240
        ACCCAGATCA ACCAAAGTGA GAAATGGAAC TACAAGAAAC ACACCAAGGA ATTTCCTACC    300
60      GACGCCTTTG GGGATATTCA GTTTGAGACA CTGGGGAAGA AGGGAAGTA TATACGTCTG    360
        TCCTGCGACA CGGACGCGGA AATCCTTTAC GAGCTGCTGA CCCAGCACTG GCACCTGAAA    420
        ACACCCAACC TGGTCATTTC TGTGACCGGG GGCGCCAAGA ACTTCGCCCT GAAGCCGCGC    480
        ATGCGCAAGA TCTTCAGCCG GCTCATCTAC ATCGCGCAGT CCAAAGGTGC TTGGATTCTC    540
        ACGGGAGGCA CCCATTATGG CCTGATGAAG TACATCGGGG AGGTGGTGAG AGATAACACC    600
65      ATCAGCAGGA GTTCAGAGGA GAATATTGTG GCCATTGGCA TAGCAGCTTG GGGCATGGTC    660
        TCCAACCGGG ACACCCTCAT CAGGAATTGC GATGCTGAGG GCTATTTTTT AGCCCAGTAC    720
        CTTATGGATG ACTTCACAAG AGATCCACTG TATATCCTGG ACAACAACCA CACACATTTG    780
        CTGCTCGTGG ACAATGGCTG TCATGGACAT CCCACTGTCG AAGCAAAGCT CCGGAATCAG    840
        CTAGAGAAGT ATATCTCTGA GCGCACTATT CAAGATTCCA ACTATGGTGG CAAGATCCCC    900
70      ATTGTGTGTT TGCCCAAGG AGGTGGAAAA GAGACTTTGA AAGCCATCAA TACCTCCATC    960
        AAAAATAAAA TTCCTTGTGT GGTGGTGGAA GGCTCGGGCC AGATCGCTGA TGTGATCGCT   1020
        AGCCTGGTGG AGGTGGAGGA TGCCCTGACA TCTTCTGCCG TCAAGGAGAA GCTGGTGCGC   1080
        TTTTTACCCC GCACGGTGTC CCGGCTGCCT GAGGAGGAGA CTGAGAGTTG GATCAAATGG   1140
        CTCAAAGAAA TTCTCGAATG TTCTCACCTA TTAACAGTTA TTAAAATGGA AGAAGCTGGG   1200
75      GATGAAATTG TGAGCAATGC CATCTCCTAC GCTCTATACA AGACCTTCAG CACCAGTGAG   1260
        CAAGACAAGG ATAACTGAA TGGGCAGCTG AAGCTTCTGC TGGAGTGGAA CCAGCTGGAC   1320
        TTAGCCAATG ATGAGATTTT CACCAATGAC CGCCGATGGG AGTCTGCTGA CCTTCAAGAA   1380
        GTCATGTTTA CGGCTCTCAT AAAGGACAGA CCCAAGTTTG TCCGCCTCTT TCTGGAGAAT   1440
        GGCTTGAACC TACGGAAGTT TCTCACCCAT GATGTCCTCA CTGAACTCTT CTCCAACCAC   1500
80      TTCAGCACGC TTGTGTACCG GAATCTGCAG ATCGCCAAGA ATTCCTATAA TGATGCCCTC   1560
```

314

```
CTCACGTTTG TCTGGAAACT GGTTGCGAAC TTCCGAAGAG GCTTCCGGAA GGAAGACAGA    1620
AATGGCCGGG ACGAGATGGA CATAGAACTC CACGACGTGT CTCCTATTAC TCGGCACCCC    1680
CTGCAAGCTC TCTTCATCTG GGCCATTCTT CAGAATAAGA AGGAACTCTC CAAAGTCATT    1740
TGGGAGCAGA CCAGGGGCTG CACTCTGGCA GCCCTGGGAG CCAGCAAGCT TCTGAAGACT    1800
CTGGCCAAAG TGAAGAACGA CATCAATGCT GCTGGGGAGT CCGAGGAGCT GGCTAATGAG    1860
TACGAGACCC GGGCTGTTGA GCTGTTCACT GAGTGTTACA GCAGCGATGA AGACTTGGCA    1920
GAACAGCTGC TGGTCTATTC CTGTGAAGCT TGGGGTGGAA GCAACTGTCT GGAGCTGGCG    1980
GTGGAGGCCA CAGACCAGCA TTTCATCGCC CAGCCTGGGG TCCAGAATTT TCTTTCTAAG    2040
CAATGGTATG GAGAGATTTC CCGAGACACC AAGAACTGGA AGATTATCCT GTGTCTGTTT    2100
ATTATACCCT TGGTGGGCTG TGGCTTTGTA TCATTTAGGA AGAAACCTGT CGACAAGCAC    2160
AAGAAGCTGC TTTGGTACTA TGTGGCGTTC TTCACCTCCC CCTTCGTGGT CTTCTCCTGG    2220
AATGTGGTCT TCTACATCGC CTTCCTCCTG CTGTTTGCCT ACGTGCTGCT CATGGATTTC    2280
CATTCGGTGC CACACCCCCC CGAGCTGGTC CTGTACTCGC TGGTCTTTGT CCTCTTCTGT    2340
GATGAAGTGA GACAGTGGTA CGTAAATGGG GTGAATTATT TTACTGACCT GTGGAATGTG    2400
ATGGACACGC TGGGGCTTTT TTACTTCATA GCAGGAATTG TATTTCGGCT CCACTCTTCT    2460
AATAAAAGCT CTTTGTATTC TGGACGAGTC ATTTTCTGTC TGGACTACAT TATTTTCACT    2520
CTAAGATTGA TCCACATTTT TACTGTAAGC AGAAACTTAG GACCCAAGAT TATAATGCTG    2580
CAGAGGATGC TGATCGATGT GTTCTTCTTC CTGTTCCTCT TTGCGGTGTG GATGGTGGCC    2640
TTTGGCGTGG CCAGGCAAGG GATCCTTAGG CAGAATGAGC AGCGCTGGAG GTGGATATTC    2700
CGTTCGGTCA TCTACGAGCC CTACCTGGCC ATGTTCGGCC AGGTGCCCAG TGACGTGGAT    2760
GGTACCACGT ATGACTTTGC CCACTGCACC TTCACTGGGA ATGAGTCCAA GCCACTGTGT    2820
GTGGAGCTGG ATGAGCACAA CCTGCCCCGG TTCCCCGAGT GGATCACCAT CCCCCTGGTG    2880
TGCATCTACA TGTTATCCAC CAACATCCTG CTGGTCAACC TGCTGGTCGC CATGTTTGGC    2940
TACACGGTGG GCACCGTCCA GGAGAACAAT GACCAGGTCT GGAAGTTCCA GAGGTACTTC    3000
CTGGTGCAGA GTACTGCAG CCGCCTCAAT ATCCCCTTCC CCTTCATCGT CTTCGCTTAC    3060
TTCTACATGG TGGTGAAGAA GTGCTTCAAG TGTTGCTGCA AGGAGAAAAA CATGGAGTCT    3120
TCTGTCTGCT GTTTCAAAAA TGAAGACAAT GAGACTCTGG CATGGGAGGG TGTCATGAAG    3180
GAAAACTACC TTGTCAAGAT CAACACAAAA GCCAACGACA CCTCAGAGGA AATGAGGCAT    3240
CGATTTAGAC AACTGGATAC AAAGCTTAAT GATCTCAAGG GTCTTCTGAA AGAGATTGCT    3300
AATAAAATCA AATGA
```

SEQ ID NO:38 PBH1 Protein sequence
Protein Accession #:   XP_017718

```
1          11         21         31         41         51
|          |          |          |          |          |
MSFRAARLSM RNRRNDTLDS TRTLYSSASR STDLSYSESD LVNFIQANFK KRECVFFTKD    60
SKATENVCKC GYAQSQHMEG TQINQSEKWN YKKHTKEFPT DAFGDIQFET LGKKGKYIRL   120
SCDTDAEILY ELLTQHWHLK TPNLVISVTG GAKNFALKPR MRKIFSRLIY IAQSKGAWIL   180
TGGTHYGLMK YIGEVVRDNT ISRSSEENIV AIGIAAWGMV SNRDTLIRNC DAEGYFLAQY   240
LMDDFTRDPL YILDNNHTHL LLVDNGCHGH PTVEAKLRNQ LEKYISERTI QDSNYGGKIP   300
IVCFAQGGGK ETLKAINTSI KNKIPCVVVE GSGQIADVIA SLVEVEDALT SSAVKEKLVR   360
FLPRTVSRLP EEETESWIKW LKEILECSHL LTVIKMEEAG DEIVSNAISY ALYKAFSTSE   420
QDKDNWNGQL KLLLEWNQLD LANDEIFTND RRWESADLQE VMFTALIKDR PKFVRLFLEN   480
GLNLRKFLTH DVLTELFSNH FSTLVYRNLQ IAKNSYNDAL LTFVWKLVAN FRRGFRKEDR   540
NGRDEMDIEL HDVSPITRHP LQALFIWAIL QNKKELSKVI WEQTRGCTLA ALGASKLLKT   600
LAKVKNDINA AGESEELANE YETRAVELFT ECYSSDEDLA EQLLVYSCEA WGGSNCLELA   660
VEATDQHFIA QPGVQNFLSK QWYGEISRDT KNWKIILCLF IIPLVGCGFV SFRKKPVDKH   720
KKLLWYVAF FTSPFVVFSW NVVFYIAFLL LFAYVLLMDF HSVPHPPELV LYSLVFVLFC   780
DEVRQWYVNG VNYFTDLWNV MDTLGLFYFI AGIVFRLHSS NKSSLYSGRV IFCLDYIIFT   840
LRLIHIFTVS RNLGPKIIML QRMLIDVFFF LFLFAVWMVA FGVARQGILR QNEQRWRWIF   900
RSVIYEPYLA MFGQVPSDVD GTTYDFAHCT FTGNESKPLC VELDEHNLPR FPEWITIPLV   960
CIYMLSTNIL LVNLLVAMFG YTVGTVQENN DQVWKFQRYF LVQEYCSRLN IPFPFIVFAY  1020
FYMVVKKCFK CCCKEKNMES SVCCFKNEDN ETLAWEGVMK ENYLVKINTK ANDTSEEMRH  1080
RFRQLDTKLN DLKGLLKEIA NKIK
```

SEQ ID NO:39 PBH3 DNA SEQUENCE

Nucleic Acid Accession #:   XM_011804
Coding sequence: 1-558 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGCCTCGCC TGTTCTTGTT CCACCTGCTA GAATTCTGTT TACTACTGAA CCAATTTTCC    60
AGAGCAGTCG CGGCCAAATG GAAGGACGAT GTTATTAAAT TATGCGGCCG CGAATTAGTT   120
CGCGCGCAGA TTGCCATTTG CGGCATGAGC ACCTGGAGCA AAAGGTCTCT GAGCCAGGAA   180
GATGCTCCTC AGACACCTAG ACCAGTGGCA GAAATTGTAC CATCCTTCAT CAACAAAGAT   240
ACAGAAACTA TAATTATCAT GTTGGAATTC ATTGCTAATT TGCCACCGGA GCTGAAGGCA   300
GCCCTATCTG AGAGGCAACC ATCATTACCA GAGCTACAGC AGTATGTACC TGCATTAAAG   360
GATTCCAATC TTAGCTTTGA AGAATTTAAG AAACTTATTC GCAATAGGCA AAGTGAAGCC   420
GCAGCAGCA ATCCTTCAGA ATTAAAATAC TTAGGCTTGG ATACTCATTC TCAAAAAAAG   480
AGACGACCCT ACGTGGCACT GTTTGAGAAA TGTTGCCTAA TTGGTTGTAC CAAAAGGTCT   540
CTTGCTAAAT ATTGCTGA
```

SEQ ID NO:40 PBH3 PROTEIN SEQUENCE

Protein Accession #:   NP_008842

```
1          11         21         31         41         51
|          |          |          |          |          |
MPRLFLFHLL EFCLLLNQFS RAVAAKWKDD VIKLCGRELV RAQIAICGMS TWSKRSLSQE    60
```

315

```
DAPQTPRPVA EIVPSFINKD TETIIIMLEF IANLPPELKA ALSERQPSLP ELQQYVPALK   120
DSNLSFEEFK KLIRNRQSEA ADSNPSELKY LGLDTHSQKK RRPYVALFEK CCLIGCTKRS   180
LAKYC
```

SEQ ID NO:41 PBH5 DNA SEQUENCE

Nucleic Acid Accession #: NM_005845
Coding sequence: 1-3978 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGCTGCCCG TGTACCAGGA GGTGAAGCCC AACCCGCTGC AGGACGCGAA CCTCTGCTCA   60
CGCGTGTTCT TCTGGTGGCT CAATCCCTTG TTTAAAATTG GCCATAAACG GAGATTAGAG   120
GAAGATGATA TGTATTCAGT GCTGCCAGAA GACCGCTCAC AGCACCTTGG AGAGGAGTTG   180
CAAGGGTTCT GGGATAAAGA AGTTTTAAGA GCTGAGAATG ACGCACAGAA GCCTTCTTTA   240
ACAAGAGCAA TCATAAAGTG TTACTGGAAA TCTTATTTAG TTTTGGGAAT TTTTACGTTA   300
ATTGAGGAAA GTGCCAAAGT AATCCAGCCC ATATTTTTGG GAAAAATTAT TAATTATTTT   360
GAAAATTATG ATCCCATGGA TTCTGTGGCT TTGAACACAG CGTACGCCTA TGCCACGGTG   420
CTGACTTTTT GCACGCTCAT TTTGGCTATA CTGCATCACT TATATTTTTA TCACGTTCAG   480
TGTGCTGGGA TGAGGTTACG AGTAGCCATG TGCCATATGA TTTATCGGAA GGCACTTCGT   540
CTTAGTAACA TGGCCATGGG GAAGACAACC ACAGGCCAGA TAGTCAATCT GCTGTCCAAT   600
GATGTGAACA AGTTTGATCA GGTGACAGTG TTCTTACACT TCCTGTGGGC AGGACCACTG   660
CAGGCGATCG CAGTGACTGC CCTACTCTGG ATGGAGATAG GAATATCGTG CCTTGCTGGG   720
ATGGCAGTTC TAATCATTCT CCTGCCCTTG CAAAGCTGTT TTGGGAAGTT GTTCTCATCA   780
CTGAGGAGTA AAACTGCAAC TTTCACGGAT GCCAGGATCA GGACCATGAA TGAAGTTATA   840
ACTGGTATAA GGATAATAAA AATGTACGCC TGGGAAAAGT CATTTTCAAA TCTTATTACC   900
AATTTGAGAA AGAAGGAGAT TTCCAAGATT CTGAGAAGTT CCTGCCTCAG GGGGATGAAT   960
TTGGCTTCGT TTTTCAGTGC AAGCAAAATC ATCGTGTTTG TGACCTTCAC CACCTACGTG   1020
CTCCTCGGCA GTGTGATCAC AGCCAGCCGC GTGTTCGTGG CAGTGACGCT GTATGGGGCT   1080
GTGCGGCTGA CGGTTACCCT CTTCTTCCCC TCAGCCATTG AGAGGGTGTC AGAGGCAATC   1140
GTCAGCATCC GAAGAATCCA GACCTTTTTG CTACTTGATG AGATATCACA GCGCAACCGT   1200
CAGCTGCCGT CAGATGGTAA AAAGATGGTG CATGTGCAGG ATTTTACTGC TTTTTGGGAT   1260
AAGGCATCAG AGACCCCAAC TCTACAAGGC CTTTCCTTTA CTGTCAGACC TGGCGAATTG   1320
TTAGCTGTGG TCGGCCCCGT GGGAGCAGGG AAGTCATCAC TGTTAAGTGC CGTGCTCGGG   1380
GAATTGGCCC CAAGTCACGG GCTGGTCAGC GTGCATGGAA GAATTGCCTA TGTGTCTCAG   1440
CAGCCCTGGG TGTTCTCGGG AACTCTGAGG AGTAATATTT TATTTGGGAA GAAATACGAA   1500
AAGGAACGAT ATGAAAAAGT CATAAAGGCT TGTGCTCTGA AAAAGGATTT ACAGCTGTTG   1560
GAGGATGGTG ATCTGACTGT GATAGGAGAT CGGGGAACCA CGCTGAGTGG AGGGCAGAAA   1620
GCACGGGTAA ACCTTGCAAG AGCAGTGTAT CAAGATGCTG ACATCTATCT CCTGGACGAT   1680
CCTCTCAGTG CAGTAGATGC GGAAGTTAGC AGACACTTGT TCGAACTGTG TATTTGTCAA   1740
ATTTTGCATG AGAAGATCAC AATTTTAGTG ACTCATCAGT TGCAGTACCT CAAAGCTGCA   1800
AGTCAGATTC TGATATTGAA AGATGGTAAA ATGGTGCAGA AGGGGACTTA CACTGAGTTC   1860
CTAAAATCTG GTATAGATTT TGGCTCCCTT TTAAAGAAGG ATAATGAGGA AAGTGAACAA   1920
CCTCCAGTTC CAGGAACTCC CACACTAAGG AATCGTACCT TCTCAGAGTC TTCGGTTTGG   1980
TCTCAACAAT CTTCTAGACC CTCCTTGAAA GATGGTGCTC TGGAGAGCCA AGATACAGAG   2040
AATGTCCCAG TTACACTATC AGAGGAGAAC CGTTCTGAAG GAAAAGTTGG TTTTTCAGGCC   2100
TATAAGAATT ACTTCAGAGC TGGTGCTCAC TGGATTGTCT TCATTTTCCT TATTCTCCTA   2160
AACACTGCAG CTCAGGTTGC CTATGTGCTT CAAGATTGGT GGCTTTCATA CTGGGCAAAC   2220
AAACAAAGTA TGCTAAATGT CACTGTAAAT GGAGGAGGAA ATGTAACCGA GAAGCTAGAT   2280
CTTAACTGGT ACTTAGGAAT TTATTCAGGT TTAACTGTAG CTACCGTTCT TTTTTGGCATA   2340
GCAAGATCTC TATTGGTATT CTACGTCCTT GTTAACTCTT CACAAACTTT GCACAACAAA   2400
ATGTTTGAGT CAATTCTGAA AGCTCCGGTA TTATTCTTTG ATAGAAATCC AATAGGAAGA   2460
ATTTTAAATC GTTTCTCCAA AGACATTGGA CACTTGGATG ATTTGCTGCC GCTGACGTTT   2520
TTAGATTTCA TCCAGACATT GCTACAAGTG GTTGGTGTGG TCTCTGTGGC TGTGGCCGTG   2580
ATTCCTTGGA TCGCAATACC CTTGGTTCCC CTTGGAATCA TTTTCATTTT TCTTCGGCGA   2640
TATTTTTTGG AAACGTCAAG AGATGTGAAG CGCCTGGAAT CTACAACTCG GAGTCCAGTG   2700
TTTTCCCACT TGTCATCTTC TCTCCAGGGG CTCTGGACCA TCCGGGCATA CAAAGCAGAA   2760
GAGAGGTGTC AGGAACTGTT TGATGCACAC CAGGATTTAC ATTCAGAGGC TTGGTTCTTG   2820
TTTTTGACAA CGTCCCGCTG GTTCGCCGTC CGTCTGGATG CCATCTGTGC CATGTTTGTC   2880
ATCATCGTTG CCTTTGGGTC CCTGATTCTG GCAAAAACTC TGGATGCCGG GCAGGTTGGT   2940
TTGGCACTGT CCTATGCCCT CACGCTCATG GGGATGTTTC AGTGGTGTGT TCGACAAAGT   3000
GCTGAGGTTG AGAATATGAT GATCTCAGTA GAAAGGGTCA TTGAATACAC AGACCTTGAA   3060
AAAGAAGCAC CTTGGGAATA TCAGAAACGC CCACCACCAG CCTGGCCCCA TGAAGGAGTG   3120
ATAATCTTTG ACAATGTGAA CTTCATGTAC AGTCCAGGTG GGCCTCTGGT ACTGAAGCAT   3180
CTGACAGCAC TCATTAAATC ACAAGAAAAG GTTGGCATTG TGGGAAGAAC CGGAGCTGGA   3240
AAAAGTTCCC TCATCTCAGC CCTTTTTAGA TTGTCAGAAC CCGAAGGTAA AATTTGGATT   3300
GATAAGATCT TGACAACTGA AATTGGACTT CACGATTTAA GGAAGAAAAT GTCAATCATA   3360
CCTCAGGAAC CTGTTTTGTT CACTGGAACA ATGAGGAAAA ACCTGGATCC CTTTAATGAG   3420
CACACGGATG AGGAACTGTG GAATGCCTTA CAAGAGGTAC AACTTAAAGA AACCATTGAA   3480
GATCTTCCTG GTAAAATGGA TACTGAATTA GCAGAATCAG GATCCAATTT TAGTGTTGGA   3540
CAAAGACAAC TGGTGTGTCCT TGCCAGGGCA ATTCTCAGGA AAAATCAGAT ATTGATTATT   3600
GATGAAGCGA CGGCAAATGT GGATCCAAGA ACTGATGAGT TAATACAAAA AAAAATCCGG   3660
GAGAAATTTG CCCACTGCAC CGTGCTAACC ATTGCACACA GATTGAACAC CATTATTGAC   3720
AGCGACAAGA TAATGGTTTT AGATTCAGGA AGACTGAAAG AATATGATGA GCCGTATGTT   3780
TTGCTGCAAA ATAAAGAGAG CCTATTTTAC AAGATGGTGC AACAACTGGG CAAGGCAGAA   3840
GCCGCTGCCC TCACTGAAAC AGCAAAACAG GTATACTTCA AAGAAAATTA TCCACATATT   3900
GGTCACACTG ACCACATGGT TACAAACACT TCCAATGGAC AGCCCTCGAC CTTAACTATT   3960
TTCGAGACAG CACTGTGA
```

SEQ ID NO:42 PBH5 PROTEIN SEQUENCE

Protein Accession #:   NP_005836

```
1          11         21         31         41         51
|          |          |          |          |          |
MLPVYQEVKP NPLQDANLCS RVFFWWLNPL FKIGHKRRLE EDDMYSVLPE DRSQHLGEEL    60
QGFWDKEVLR AENDAQKPSL TRAIIKCYWK SYLVLGIFTL IEESAKVIQP IFLGKIINYF   120
ENYDPMDSVA LNTAYAYATV LTFCTLILAI LHHLYFHVQ CAGMRLRVAM CHMIYRKALR   180
LSNMAMGKTT TGQIVNLLSN DVNKFDQVTV FLHFLWAGPL QAIAVTALLW MEIGISCLAG   240
MAVLIILLPL QSCFGKLFSS LRSKTATFTD ARIRTMNEVI TGIRIIKMYA WEKSFSNLIT   300
NLRKKEISKI LRSSCLRGMN LASFFSASKI IVFVTFTTYV LLGSVITASR VFVAVTLYGA   360
VRLTVTLFFP SAIERVSEAI VSIRRIQTFL LLDEISQRNR QLPSDGKKMV HVQDFTAFWD   420
KASETPTLQG LSFTVRPGEL LAVVGPVGAG KSSLLSAVLG ELAPSHGLVS VHGRIAYVSQ   480
QPWVFSGTLR SNILFGKKYE KERYEKVIKA CALKKDLQLL EDGDLTVIGD RGTTLSGGQK   540
ARVNLARAVY QDADIYLLDD PLSAVDAEVS RHLFELCICQ ILHEKITILV THQLQYLKAA   600
SQILILKDGK MVQKGTYTEF LKSGIDFGSL LKKDNEESEQ PPVPGTPTLR NRTFSESSVW   660
SQQSSRPSLK DGALESQDTE NVPVTLSEEN RSEGKVGFQA YKNYFRAGAH WIVFIFLILL   720
NTAAQVAYVL QDWWLSYWAN KQSMLNVTVN GGGNVTEKLD LNWYLGIYSG LTVATVLFGI   780
ARSLLVFYVL VNSSQTLHNK MFESILKAPV LFFDRNPIGR ILNRFSKDIG HLDDLLPLTF   840
LDFIQTLLQV VGVVSVAVAV IPWIAIPLVP LGIIFIFLRR YFLETSRDVK RLESTTRSPV   900
FSHLSSSLQG LWTIRAYKAE ERCQELFDAH QDLHSEAWFL FLTTSRWFAV RLDAICAMFV   960
IIVAFGSLIL AKTLDAGQVG LALSYALTLM GMFQWCVRQS AEVENMMISV ERVIEYTDLE  1020
KEAPWEYQKR PPPAWPHEGV IIFDNVNFMY SPGGPLVLKH LTALIKSQEK VGIVGRTGAG  1080
KSSLISALFR LSEPEGKIWI DKILTTEIGL HDLRKKMSII PQEPVLFTGT MRKNLDPFNE  1140
HTDEELWNAL QEVQLKETIE DLPGKMDTEL AESGSNFSVG QRQLVCLARA ILRKNQILII  1200
DEATANVDPR TDELIQKKIR EKFAHCTVLT IAHRLNTIID SDKIMVLDSG RLKEYDEPYV  1260
LLQNKESLFY KMVQQLGKAE AAALTETAKQ VYFKRNYPHI GHTDHMVTNT SNGQPSTLTI  1320
FETAL
```

SEQ ID NO:43 PBQ7 DNA SEQUENCE

Nucleic Acid Accession #: NM_021233
Coding sequence: 34-1119 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGGGGAAAG TGTCCTGCTG TGGCATGAAA TAAATGAAAC AGAAAATGAT GGCAAGACTG    60
CTAAGAACAT CCTTTGCTTT GCTCTTCCTT GGCCTCTTTG GGGTGCTGGG GGCAGCAACA   120
ATTTCATGCA GAAATGAAGA AGGGAAAGCT GTGGACTGGT TTACTTTTTA TAAGTTACCT   180
AAAAGACAAA ACAAGGAAAG TGGAGAGACT GGGTTAGAGT ACCTGTACCT AGACTCTACA   240
ACTAGAAGCT GGAGGAAGAG TGAGCAACTA ATGAATGACA CCAAGAGTGT TTTGGGAAGG   300
ACATTACAAC AGCTATATGA AGCATATGCC TCTAAGAGTA ACAACACAGC CTATCTAATA   360
TACAATGATG GAGTCCCTAA ACCTGTGAAT TACAGTAGAA AGTATGGACA CACCAAAGGT   420
TTACTGCTGT GGAACAGAGT TCAAGGGTTC TGGCTGATTC ATTCCATCCC TCAGTTTCCT   480
CCAATTCCGG AAGAAGGCTA TGATTATCCA CCCACAGGGA GACGAAATGG ACAAAGTGGC   540
ATCTGCATAA CTTTCAAGTA CAACCAGTAT GAGGCAATAG ATTCTCAGCT CTTGGTCTGC   600
AACCCCAACG TCTATAGCTG CTCCATCCCA GCCACCTTTC ACCAGGAGCT CATTCACATG   660
CCCCAGCTGT GCACCAGGGC CAGCTCATCA GAGATTCCTG GCAGGCTCCT CACCACACTT   720
CAGTCGGCCC AGGGACAAAA ATTCCTCCAT TTTGCAAAGT CGGATTCTTT TCTTGACGAC   780
ATCTTTGCAG CCTGGATGGC TCAACGGCTG AAGACACACT TGTTAACAGA AACCTGGCAG   840
CGAAAAAGAC AAGAGCTTCC TTCAAACTGC TCCCTTCCTT ACCATGTCTA CAATATAAAA   900
GCAATTAAAT TATCACGACA CTCTTATTTC AGTTCTTATC AAGATCACGC CAAGTGGTGT   960
ATTTCCCAAA AGGGCACCAA AAATCGCTGG ACATGTATTG GAGACCTAAA TCGGAGTCCA  1020
CACCAAGCCT TCAGAAGTGG AGGATTCATT TGTACCCAGA ATTGGCAAAT TTACCAAGCA  1080
TTTCAAGGAT TAGTATTATA CTATGAAAGC TGTAAGTAA CTTGGTGAAA GGACACAGGT
```

SEQ ID NO:44 PBQ7 Protein sequence

Protein Accession #:   NP_067056

```
1          11         21         31         41         51
|          |          |          |          |          |
MMARLLRTSF ALLFLGLFGV LGAATISCRN EEGKAVDWFT FYKLPKRQNK ESGETGLEYL    60
YLDSTTRSWR KSEQLMNDTK SVLGRTLQQL YEAYASKSNN TAYLIYNDGV PKPVNYSRKY   120
GHTKGLLLWN RVQGFWLIHS IPQFPPIPEE GYDYPPTGRR NGQSGICITF KYNQYEAIDS   180
QLLVCNPNVY SCSIPATFHQ ELIHMPQLCT RASSSEIPGR LLTTLQSAQG QKFLHFAKSD   240
SFLDDIFAAW MAQRLKTHLL TETWQRKRQE LPSNCSLPYH VYNIKAIKLS RHSYFSSYQD   300
HAKWCISQKG TKNRWTCIGD LNRSPHQAFR SGGFICTQNW QIYQAFQGLV LYYESCK
```

SEQ ID NO:45 PCQ8 DNA SEQUENCE

Nucleic Acid Accession #: XM_030453
Coding sequence: 89-1273 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
CGGTGCCCTG GGGTGGAATA TCCCCTACGA ATTTAACCAA GCGGACTTTA ATGCCACTGT    60
GCAGTTCATC CAAAACCACT TGGATGACAT GGATGTCAAA AAGGGTGTCT CCTGGACCAC   120
CATCCGCTAC ATGATAGGAG AGATTCAATA TGGAGGCAGA GTCACTGACG ACTATGATAA   180
GAGATTGTTG AACACATTTG CTAAGGTTTG GTTCAGTGAA AATATGTTTG GACCAGATTT   240
CAGTTTTTAC CAAGGATACA ATATTCCAAA ATGCAGCACA GTGGATAACT ATCTTCAGTA   300
TATCCAGAGT TTGCCTGCCT ATGACAGCCC TGAGGTGTTT GGGCTGCACC CCAATGCTGA   360
```

```
CATCACCTAC CAGAGCAAGC TGGCCAAGGA CGTGCTGGAC ACCATCCTAG GCATCCAACC    420
CAAGGACACC TCTGGTGGAG GGGATGAGAC CCGGGAGGCG GTGGTGGCCC GGCTGGCTGA    480
TGATATGCTG GAGAAGCTGC CCCCAGACTA TGTCCCCTTT GAAGTAAAAG AGAGGCTGCA    540
GAAGATGGGG CCATTCCAGC CTATGAACAT TTTCCTCAGG CAGGAAATAG ACAGAATGCA    600
AAGGGTACTC AGCCTTGTCC GCAGCACCCT CACTGAGCTG AAACTTGCTA TTGATGGCAC    660
CATCATCATG AGCGAAAATC TGCAAGATGC ATTGGATTGC ATGTTTGATG CTAGAATCCC    720
TGCTTGGTGG AAAAAAGCTT CTTGGGTTTT TAGTACACTG GGTTTCTGGT TTACTGAACT    780
TATAGAAAGA AACAGCCAGT TTACCTCGTG GGTTTTCAAT GGCCGACCTC ACTGCTTTTG    840
GATGACGGGT TTTTTTAACC CCCAGGGATT TTTAACTGCA ATGCGACAGG AAATAACTCG    900
GGCCAACAAA GGCTGGGCTC TGGACAATAT GGTGCTTTGC AATGAAGTCA CCAAATGGAT    960
GAAGGACGAC ATTTCTACCC CTCCCACAGA GGGTGTCTAT GTCTATGGCT TATATCTTGA   1020
AGGTGCTGGC TGGGACAAGA GGAACATGAA ACTCATTGAA TCAAAGCCAA AAGTGCTCTT   1080
TGAGTTGATG CCTGTCATAA GGATTTATGC AGAAAACAAT ACTTTACGAG ATCCTCGGTT   1140
TTACTCCTGT CCCATCTATA AGAAGCCAGT TCGAACGGAC TTGAACTACA TTGCCGCTGT   1200
GGATCTCAGG ACAGCCCAGA CCCCTGAACA CTGGGTGCTC CGTGGGGTTG CCCTTCTGTG   1260
TGATGTCAAG TAACATGTGG GGAGTGTCCC CACCCAATGC TTTGGAAAAT GCAAGATCTA   1320
AATTATTGTA ACCTTTATTT CTGTATGACT GCTGGACAGT GTATGTTAGG TCGTTTATGC   1380
AATTAATGAG CTGCATAGGT TTTCCCCACT CCTTAATTGG ATGCTTATAT TTTACTTGTT   1440
TCATCATTAG TGACCAATGT CTGAGTTTGT TGAAAATGTT ATTTAGTGAT ATAAAAGTAA   1500
ATTTACAGCA TCCTAATGAA GTGTGGCCCT CAAATCCACA GTAGTATATT TTCTTCTTAC   1560
TTCGCTCCGA AGACTGACTG TGATTATAAC AGCAAATATA TTTGCATGTG GACAAAGATT   1620
AGATGGCAAG ATAGAAAAAT AAGAACAGAT GTGATAGCAA GAATTATAGT TGGCTTGAAA   1680
AAATGTGATG ATCAGGAGAA AAAATAAAAA AAGGGTAGAA ATATTAGACG GTGCGTAGGG   1740
ACTTTCTATG GACTTTTATT AATTAGGAAA CATTATCAAA GGAACTTTTC ACGTATTTTT   1800
CTTTAAATTC TGGTTAGATG TTATTAATAA TTCTTCATCT AACCTACTGA CTAGAAAATA   1860
TAGTCAGTAC TAAATTAGAA TTGTGGTTTA TAAACTTTTG GTTAGCTCTG GATCTGTATA   1920
ACTGCATTTT TTTGGATAAA CAGTTTTTGG TAGGTGGATA CCGGGAGACA AGTGTGGGTC   1980
CCTCTCACTG GGCTTCATTC TGTGGACCAG GATCATTATT TCATGCTCAT GATCATGAGA   2040
GTTAGGACTG AGTGGCTCCT GTGACTCCCA CCATCTTAGA TGATACTGTT TTCTTGTGAG   2100
TTCTTTCTTT TGGTGTGGAT TAGTATATCA GTTGATTTGT GTGAATTGTG GTGAAACAAT   2160
CATTTCATTT TGAAAAGCAA GTAATGAAAA TGTCAGCATC ATAGGAATTA ATAAAATGTT   2220
TTTACTAAAA AAAAAAAAAA AAA
```

SEQ ID NO:46 PCQ8 Protein sequence
Protein Accession #:   BAB15543

```
1          11         21         31         41         51
|          |          |          |          |          |
MDVKKGVSWT TIRYMIGEIQ YGGRVTDDYD KRLLNTFAKV WFSENMFGPD FSFYQGYNIP    60
KCSTVDNYLQ YIQSLPAYDS PEVFGLHPNA DITYQSKLAK DVLDTILGIQ PKDTSGGGDE   120
TREAVVARLA DDMLEKLPPD YVPFEVKERL QKMGPFQPMN IFLRQEIDRM QRVLSLVRST   180
LTELKLAIDG TIIMSENLQD ALDCMFDARI PAWWKKASWV FSTLGFWFTE LIERNSQFTS   240
WVFNGRPHCF WMTGFFNPQG FLTAMRQEIT RANKGWALDN MVLCNEVTKW MKDDISTPPT   300
EGVYVYGLYL EGAGWDKRNM KLIESKPKVL FELMPVIRIY AENNTLRDPR FYSCPIYKKP   360
VRTDLNYIAA VDLRTAQTPE HWVLRGVALL CDVK
```

SEQ ID NO:47 PDG5 DNA SEQUENCE

Nucleic Acid Accession #:   AB033036
Coding sequence: 68-3349 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
GGAGCAGCCT ACAACTTCAC AACCAGAAAC CACTACCCCT CAGGGGTTGC TTTCAGATAA    60
AGATGACATG GGAAGGAGAA ATGCTGGCAT AGATTTCGGA TCCAGAAAAG CATCAGCAGC   120
ACAGCCCATA CCTGAAAACA TGGACAATTC CATGGTTAGT GATCCACAAC CATACCATGA   180
AGATGCAGCT TCTGGAGCTG AGAAGACAGA AGCCAGAGCT TCTCTCTCAC TGATGGTGGA   240
AAGCCTTTCT ACAAACCAAG AGGAGGCCAT TCTCTCAGTA GCAGCAGAGG CTCAGGTGTT   300
TATGAATCCT TCTCATATCC AGTTAGAAGA TCAAGAAGCT TTCAGCTTTG ATTTACAAAA   360
GGCCCAATCC AAAATGGAGT CAGCCCAGGA TGTTCAAACT ATCTGCAAAG AAAAGCCTTC   420
TGGAAATGTT CACCAGACCT TTACAGCAAG TGTTTTGGGT ATGACAAGTA CTACAGCCAA   480
AGGAGATGTT TATGCCAAGA CTCTGCCTCC CAGAAGCCTT TTTCAGTCCT CAAGGAAGCC   540
TGATGCTGAA GAAGTCTCCT CAGATTCAGA GAATATTCCT GAGGAGGGGG ATGGTTCTGA   600
AGAACTGGCT CATGGTCACT CTTCCCAGTC CTTGGGGAAG TTTGAAGATG AACAAGAAGT   660
CTTCTCAGAA TCAAAAAGTT TTGTTGAGGA CTTGAGCAGC TCTGAGGAGG AGCTGGACCT   720
CAGATGCCTC TCCCAGGCTT TAGAGGAGCC TGAAGATGCA GAAGTCTTCA CAGAATCAAG   780
CAGTTATGTT GAAAAGTACA ACACTTCTGA TGATTGCAGC AGCTCAGAGG AAGACCTGCC   840
TCTCAGACAC CCTGCTCAGG CCTTGGGAAA GCCCAAAAAC CAACAAGAAG TCTCCTCTGC   900
TTCAAATAAT ACTCCTGAAG AGCAGAATGA TTTTATGCAG CAGCTGCCTT CCAGATGCCC   960
TTCTCAGCCC ATTATGAATC CTACTGTTCA GCAACAAGTC CCCACCAGTT CAGTGGGCAC   1020
TTCTATAAAA CAGAGCGATT CCGTGGAGCC AATCCCTCCA AGACACCCTT TCCAGCCATG   1080
GGTGAACCCT AAAGTGGAGC AAGAAGTTTC TCATCTCCA AAGAGCATGG CTGTTGAAGA   1140
GAGCATTTCT ATGAAGCCTC TGCCTCCTAA ACTTCTTTGC CAGCCCTTGA TGAATCCTAA   1200
AGTTCAACAA AACATGTTCT CAGGTTCAGA GGACATTGCT GTTGAGAGAG TCATTTCTGT   1260
GGAGCCACTA CTCCCCAGAT ATTCTCCTCA GTCCTTGACA GATCCTGCCTC TCCGGCAAAT   1320
CTCAGAAAGC ACAGCTGTTG AGGAAGGCAC TTATGTGGAA CCGCTGCCTC CCAGATGCCT   1380
TTCCCAGCCC TCGGAGAGGC CTAAGTTCCT GGACTCAATG AGTACTTCTG CAGAATGGAG   1440
CAGTCCTGTG GCACCAACAC CTTCCAAATA CACTTCCCCG CCATGGGTGA CCCCTAAATT   1500
TGAGGAACTG TATCAACTCT CTGCACATCC AGAAAGCACT ACTGTTGAAG AGGACATTTC   1560
TAAGGAGCAG CTGCTTCCCA GACATCTTTC CCAGTTGACT GTGGGAAATA AAGTCCAGCA   1620
ACTGTCCTCA AATTTCGAGC GGGCTGCTAT TGAGGCAGAC ATTTCTGGGA GTCCATTGCC   1680
```

318

```
        TCCCCAATAT GCTACCCAGT TCTTAAAGAG GTCTAAAGTT CAGGAAATGA CCTCACGACT  1740
        AGAGAAAATG GCTGTTGAAG GCACTTCTAA CAAATCACCG ATTCCCAGGC GTCCGACCCA  1800
        GTCATTCGTG AAATTTATGG CACAGCAAAT CTTTTCAGAG AGCTCTGCTC TTAAGAGGGG  1860
        CAGTGATGTG GCACCTCTGC CTCCCAATCT TCCTTCCAAA TCTTTATCAA AGCCTGAAGT  1920
5       CAAGCACCAA GTTTTCTCAG ATTCAGGGAG TGCTAATCCT AAGGGAGGCA TTTCTTCAAA  1980
        GATGCTACCT ATGAAGCACC CTTTACAGTC CTTGGGGAGG CCTGAAGACC CACAGAAAGT  2040
        TTTCTCTTAT TCAGAGAGAG CTCCTGGGAA GTGCAGCAGT TTTAAAGAGC AGCTGTCTCC  2100
        CAGGCAGCTT TCCCAGGCCT TGAGGAAACC TGAGTATGAG CAAAAAGTCT CCCCTGTTTC  2160
        TGCCAGTTCT CCTAAAGAGT GGAGGAATTC TAAAAAGCAG CTGCCTCCCA AACATTCTTC  2220
10      CCAAGCCTCA GATAGGTCTA AATTCCAGCC ACAGATGTCA TCAAAGGGCC CAGTGAATGT  2280
        ACCTGTAAAG CAGAGCAGCG GTGAGAAGCA CCTGCCTTCA AGTAGTCCTT TCCAGCAACA  2340
        GGTTCATTCA AGTTCTGTGA ATGCTGCTGC TAGGCGATCT GTTTTTGAGA GCAATTCTGA  2400
        CAATTGGTTC CTAGGAAGAG ATGAAGCTTT TGCAATCAAA ACCAAGAAAT TCAGCCAAGG  2460
        TTCCAAAAAC CCCATAAAGA GCATTCCAGC CCCTGCTACC AAACCTGGGA AGTTCACCAT  2520
15      TGCTCCTGTC AGGCAAACAT CCACTTCTGG GGGCATTTAC TCTAAGAAAG AAGATCTTGA  2580
        GAGTGGTGAT GGTAATAATA ACCAGCATGC AAACCTATCC AATCAGGATG ATGTTGAAAA  2640
        GCTTTTTGGA GTTCGACTGA AAAGAGCCCC TCCCTCGCAG AAGTATAAGA GTGAGAAACA  2700
        AGATAACTTC ACCCAGCTTG CTTCAGTGCC CTCGGGCCCA ATTTCATCCT CTGTAGGCAG  2760
        GGGACATAAA ATCAGAAGCA CTTCCCAGGG GCTCCTGGAT GCTGCAGGGA ACCTCACCAA  2820
20      AATATCTTAC GTTGCAGATA AGCAACAGAG CAGGCCCAAA TCTGAAAGCA TGGCCAAGAA  2880
        GCAACCTGCT TGCAAGACCC CAGGAAAGCC TGCTGGTCAA CAGTCAGATT ATGCTGTCTC  2940
        AGAGCCGGTT TGGATAACTA TGGCAAAGCA GAAGCAGAAG AGTTTCAAGG CCCACATTTC  3000
        TGTGAAAGAG CTGAAAACTA AGAGCAATGC TGGGAGCCGAT GCTGAGACTA AGGAGCCTAA  3060
        ATATGAGGGA GCTGGCTCTG CAAATGAAAA CCAACCTAAA AAGATGTTCA CTTCCAGTGT  3120
25      CCATAAACAG GAGAAGACAG CACAGATGAA GCCACCTAAG CCTACAAAAT CAGTTGGATT  3180
        TGAAGCTCAG AAGATACTGC AAGTTCCTGC CATGGAAAAA GAAACCAAAC GATCTTCAAC  3240
        TCTCCCAGCC AAGTTCCAGA ACCCAGTTGA GCCAATTGAG CCTGTCTGGT TCTCACTGGC  3300
        CAGGAAGAAA GCCAAAGCAT GGAGCCACAT GGCAGAAATC ACGCAATAAA GAGCTCTTGT  3360
        GTGGAGCATC AGCATTTATT TTATTTAGTT TTTTTTTTTT TTTTTTTTTT GAGACAGAGT  3420
30      CTCGCTCTGT TACCCAGATT GGAGTGCAGT GGCGCGATCT CCGCTCACTG CAAGCTCCGC  3480
        CTCCCGGGTT CACGCCACTC TCCCGCCTCA GTCTCCCGAC TAGCTGGGAC TACAGGCGCC  3540
        CGCCATCACG CCCGGCTAAT TTTGTTTTCG TATTTTTAGT AGAGACGGGG TTTCACCATG  3600
        TTGGCCAGGA TGGTCTTGAT CTCCTGACCT CGTGATCCGC CCGCCTCAGC CTCCCAAAAG  3660
        CTGGGATTAC AGGCGTGAGC CACCGCGCCC GGCCAAGCAT CAGCGTTTTA AATGATAATT  3720
35      GCTAATAGCT GTATTAATTC TATGTAGTGA TCTTTTTACT GTGACCACTT GTATTAAGCA  3780
        AAATAAGTAT TAAGCAAACT AAGAATTTAT TAAGCAAAAT AAGAATTTAT TAAGCAAAAT  3840
        AGCCTTAGAA ATGCAAATTA AAACATAATT ATTTGAATGA AATAAATGCC ATGAATGCTT  3900
        AACCTTCCAC GTAGTCACTG CCAGCACCCA GAAACCCAGC ATTTCCTCTA TTAAAACTAT  3960
        CGAAAACATT TGCACTGCTG TAAAATTGCA AAATCTTTAA CTTTGGACAA TGTGCTTTAG  4020
40      AAGGGAGAAA GCAAAAACAT TTTGTTGGAG CAACTAGAAA ATTGTCATTT CCCTCAACCA  4080
        AATAAAGTAA TTCTAATGGA AACATTCAGA TGATTTGACC TAAAGATTGG CCTTTAGGTT  4140
        TTATGAGCCT AGATAGATGC CGCAATTATT TGGTTGTTGC TCTAAGCTTT GCAAGGGATC  4200
        CTAAAAGAGG CGGTGGAAGT GAAAATTCTG GGTCTCCAAG AAAATTTCTG CACAGCCAGT  4260
        TCTCCAATCA GCCTATCACC CCTTGAAACA TCTTCCCTGT GTCCCTGGGG GCCCCTGATG  4320
45      CTTTCTCCTT GGGTGATAGT AACATGCAGA GCACTTACAC AAAGCTCCCC CTTTGGACAT  4380
        ACCCCACGTC GACCTGTCAC AGGCCTGGCT GTAGCGAGCA CCTCCCTATG ACGCAGAATG  4440
        CTTCTTGGGA ATTATCTTAC TCCTCTGGAG GGTTAGTCCA TCAATGTTTT GCTTCTTGTC  4500
        CCAATACTAC TGTGACCCTC TCTGATCGCA CAGAAATCAC TGCCTATCAC ATATATCCTG  4560
        TTAAGCACTG AAGACCCTAT TGAAATTAGA GTTCTACAGA TGCCAAAAGC TGTACTTTCC  4620
50      ATCAGGCAGA TGGCAAGCTT ACTGCCTTGA TGCACATCTG GAGCCACTGG AGCTCCTTCC  4680
        TCTCTGGTTC CAGCATTAAG GTGGAGAACT CCATGTAGCT TCTTGTCCTT TCCCCTCAGC  4740
        TGTCTTTGCT TCACAAGGTT TTAGCCCAAA GCAAGAGTGC AATCCCAAAG CCACAGAGAA  4800
        ATGAACTTTC CGCTACCTGG AAGCTTTAAG TGAGTAAATC AGCTTTTCCC CTCTCATTCC  4860
        TAGAGGCACA CACCTCAAAA GTTACTAGGC TGGAGAGACC CTACCTTCCA GTGACCCACT  4920
55      CATCCCCCAG CCACGGAGAA GAGGGAAGAC CAAAAAGGGA GAGTGAGAAA GAGGATGAGA  4980
        GGGATGGTCA GCTGTGAGGG GAGGGGGCAA GTGGCCCAGC AAATGTTGAT GCCTCCCTTC  5040
        CCATCTTGCC ACACGGTCTT TTTCTTTTGT AGCACAGCCT CCATTAATAA CTCCTCGGCT  5100
        GAGGATGAAG ATGTAGGCAC CTTTACCCCC AGAGCCAGTT CCTTAATTGG CTGGCTTTCT  5160
        GAGATGCAGA CCACCCTAGA ATCTCATCTA GGTTCACTAG AAGTTAGTTA AATCTTCCTT  5220
60      TCTCTGTCTT TCTCTTCATT CCATCCCCCA AACCCACCAA ACACTAAGGG AGAGCTCCCT  5280
        TTGGATGTCT GGGCAGTAAA CCTAGCTCAT TTTTCTAGGA GACCCAGAAG TGACTTCTGA  5340
        GTAGTTATCA CTGTGTCTGC CTCTGTTACA CTGTGCTGCT TTGCTTAAAC AGAAATGCAG  5400
        GCCTGGACAT CTGACTGTGC CTTTATATTC TGAGTGGGGT GCTGCCCCAT GCAAAAAAAT  5460
        CCAGAGAGGT AGTGAGGTGT CAGAGCTAAA CACTTGGTGC TGGGTTTTGT TGATGCTGGT  5520
65      ATAATGTGAC ACAGTACAAT TACATGCTAA ATTTTGCATT TTCTCTATAT AACATCTATT  5580
        TTTCCTGATA CTGTGCCTTT GCCATTTTGA TAATGCTATT TTGATTGAGT GAATTTTATT  5640
        TCCTTTGTAT TCCCATAGTG AACAATATAT TAAGGTAGAT GCCCTTTATC TGGGTACTCC  5700
        TGGTAGATTA GCTGTTACAC CTCCCTTCCC TTTTTTACAG TGAACCTGTA TTCAGTTATT  5760
        GTCACTCTGA GAACTCTCCA ATAACAATTT CTTTTCCACA GTTAACAACA CAGCTGTTAC  5820
70      ACCTCCCTTC CTTTTTTTCA CAGTGAACCT GTATTCAGCT ATTCTCACTC TGAGAACTCT  5880
        CCAATAACAA TTTCTTTTCC ACAGTTAACA ACAAAGTTCT GTTTTTAAAT GAAGAGATTA  5940
        AGTTCTTTTT AAATGCCTAA AGGCATATTC TGACAACTTT TCTACTTCTT TAACTTTTTT  6000
        GATTTAAGAT ATATGCAAAG CAAATAAATT CAATAAAGCC T
```

SEQ ID NO:48 PDG5 Protein sequence
Protein Accession #:   BAA86524

```
1          11         21         31         41         51
|          |          |          |          |          |
EQPTTSQPET TTPQGLLSDK DDMGRRNAGI DFGSRKASAA QPIPENMDNS MVSDPQPYHE     60
```

319

```
      DAASGAEKTE ARASLSLMVE SLSTTQEEAI LSVAAEAQVF MNPSHIQLED QEAFSFDLQK    120
      AQSKMESAQD VQTICKEKPS GNVHQTFTAS VLGMTSTTAK GDVYAKTLPP RSLFQSSRKP    180
      DAEEVSSDSE NIPEEGDGSE ELAHGHSSQS LGKFEDEQEV FSESKSFVED LSSSEEELDL    240
   5  RCLSQALEEP EDAEVFTESS SYVEKYNTSD DCSSSEEDLP LRHPAQALGK PKNQQEVSSA    300
      SNNTPEEQND FMQQLPSRCP SQPIMNPTVQ QQVPTSSVGT SIKQSDSVEP IPPRHPFQPW    360
      VNPKVEQEVS SSPKSMAVEE SISMKPLPPK LLCQPLMNPK VQQNMFSGSE DIAVERVISV    420
      EPLLPRYSPQ SLTDPQIRQI SESTAVEEGT YVEPLPPRCL SQPSERPKFL DSMSTSAEWS    480
      SPVAPTPSKY TSPPWVTPKF EELYQLSAHP ESTTVEEDIS KEQLLPRHLS QLTVGNKVQQ    540
  10  LSSNFERAAI EADISGSPLP PQYATQFLKR SKVQEMTSRL EKMAVEGTSN KSPIPRRPTQ    600
      SFVKFMAQQI FSESSALKRG SDVAPLPPNL PSKSLSKPEV KHQVFSDSGS ANPKGGISSK    660
      MLPMKHPLQS LGRPEDPQKV FSYSERAPGK CSSFKEQLSP RQLSQALRKP EYEQKVSPVS    720
      ASSPKEWRNS KKQLPPKHSS QASDRSKFQP QMSSKGPVNV PVKQSSGEKH LPSSSPFQQQ    780
  15  VHSSSVNAAA RRSVFESNSD NWFLGRDEAF AIKTKKFSQG SKNPIKSIPA PATKPGKFTI    840
      APVRQTSTSG GIYSKKEDLE SGDGNNNQHA NLSNQDDVEK LFGVRLKRAP PSQKYKSEKQ    900
      DNFTQLASVP SGPISSSVGR GHKIRSTSQG LLDAAGNLTK ISYVADKQQS RPKSESMAKK    960
      QPACKTPGKP AGQQSDYAVS EPVWITMAKQ KQKSFKAHIS VKELKTKSNA GADAETKEPK   1020
      YEGAGSANEN QPKKMFTSSV HKQEKTAQMK PPKPTKSVGF EAQKILQVPA MEKETKRSST   1080
  20  LPAKFQNPVE PIEPVWFSLA RKKAKAWSHM AEITQ
```

SEQ ID NO:49 PAB7 DNA SEQUENCE

Nucleic Acid Accession #: D87742

Coding sequence: 208-3582 (underlined sequences correspond to start and stop codons)

```
      1          11         21         31         41         51
      |          |          |          |          |          |
      GCTTTCCTTT CTAAAGTAGA AGAGGATGAT TATCCCTCTG AAGAACTACT AGAGGATGAA     60
      AACGCTATAA ATGCAAAACG GTCTAAAGAA AAAAACCCTG GGAATCAGGG CAGGCAGTTT    120
      GATGTTAATC TGCAAGTCCC TGACAGAGCA GTTTTAGGGA CCATTCATCC AGATCCAGAA    180
      ATTGAAGAAA GCAAGCAAGA AACTAGTATG ATTTTGGATA GTGAAAAAAC AAGTGAGACT    240
      GCTGCCAAAG GGGTCAACAC AGGAGGCAGG GAACCAAATA CAATGGTGGA AAAAGAACGC    300
      CCTCTGGCAG ATAAGAAAGC ACAGAGACCA TTTGAACGAA GTGACTTTTC TGACAGCATA    360
      AAAATTCAGA CTCCAGAATT AGGTGAAGTG TTTCAGAATA AAGATTCTGA TTATCTGAAG    420
      AACGACAACC CTGAGGAACA TCTGAAGACC TCAGGGCTTG CAGGGGAGCC TGAGGGAGAA    480
      CTCTCAAAAG AGGACCATGG GAACACAGAG AAGTACATGG GCACAGAAGA CCAGGGGTCT    540
      GCTGCTGCAG AACCTGAAGA TGACTCGTTC CACTGGACTC CACATACAAG TGTAGAGCCA    600
      GGGCATAGTG ACAAGAGGGA GGACTTACTT ATCATAAGCA GCTTCTTTAA AGAACAACAG    660
      TCTTTGCAGC GGTTCCAGAA GTACTTTAAT GTCCATGAGC TGGAAGCCTT GCTACAAGAA    720
      ATGTCATCAA AACTGAAGTC AGCGCAGCGA GAGAGCCTGC CCTATAATAT GGAAAAAGTC    780
      CTAGATAAGG TCTTCCGTGC TTCTGAGTCA CAAATTCTGA GCATAGCAGA AAAAAATGCTT   840
      GATACTCGTG TGGCTGAAAA TAGAGATCTG GGAATGAACG AAAATAACAT ATTTGAAGAG    900
      GCTGCAGTGC TTGATGACAT TCAAGACCTC ATCTATTTTG TCAGGTACAA GCACTCCACA    960
      GCAGAGGAGA CAGCCACACT GGTGATGGCA CCACCTCTAG AGGAAGGCTT GGGTGGAGCA   1020
      ATGGAAGAGA TGCAACCACT GCATGAAGAT AATTTCTCAC GAGAGAAGAC AGCAGAACTT   1080
      AATGTGCAGG TTCCTGAAGA ACCCACCCAC TTGGACCAAC GTGTGATTGG GGACACTCAT   1140
      GCCTCAGAAG TGTCACAGAA GCCAAATACT GAGAAAGACC TGGACCCAGG GCCAGTTACA   1200
      ACAGAAGACA CTCCTATGGA TGCTATTGAT GCAAACAAGC AACCAGAGAC AGCCGCCGAA   1260
      GAGCCGGCAA GTGTCACACC TTTGGAAAAC GCAATCCTTC TAATATATTC ATTCATGTTT   1320
      TATTTAACTA AGTCGCTAGT TGCTACATTG CCTGATGATG TTCAGCCCTGG GCCTGATTTT   1380
      TATGGACTGC CATGGAAACC TGTATTTATC ACTGCCTTCT TGGGAATTGC TTCGTTTGCC   1440
      ATTTTCTTAT GGAGAACTGT CCTTGTTGTG AAGGATAGAG TATATCAAGT CACGGAACAG   1500
      CAAATTTCTG AGAAGTTGAA GACTATCATG AAAGAAAATA CAGAACTTGT ACAAAAATTG   1560
      TCAAATTATG AACAGAAGAT CAAGGAATCA AAGAAACATG TTCAGGAAAC CAGGAAACAA   1620
      AATATGATTC TCTCTGATGA AGCAATTAAA TATAAGGATA AAATCAAGAC ACTTGAAAAA   1680
      AATCAGGAAA TTCTGGATGA CACAGCTAAA AATCTTCGTG TTATGCTAGA ATCTGAGAGA   1740
      GAACAGAATG TCAAGAATCA GGACTTGATA TCAGAAAACA AGAAATCTAT AGAGAAGTTA   1800
      AAGGATGTTA TTTCAATGAA TGCCTCAGAA TTTTCAGAGG TTCAGATTGC ACTTAATGAA   1860
      GCTAAGCTTA GTGAAGAGAA GGTGAAGTCT GAATGCCATC GGGTTCAAGA AGAAAATGCT   1920
      AGGCTTAAGA AGAAAAAAGA GCAGTTGCAG CAGGAAATCG AAGACTGGAG TAAATTACAT   1980
      GCTGAGCTCA GTGAGCAAAT CAAATCATTT GAGAAGTCTC AGAAAGATTT GGAAGTAGCT   2040
      CTTACTCACA AGGATGATAA TATTAATGCT TTGACTAACT GCATTACACA GTTGAATCTG   2100
      TTAGAGTGTG AATCTGAATC TGAGGGTCAA AATAAAGGTG GAAATGATTC AGATGAATTA   2160
      GCAAATGGAG AAGTGGGAGG TGACCGGAAT GAGAAGATGA AAAATCAAAT TAAGCAGATG   2220
      ATGGATGTCT CTCGGACACA GACTGCAATA TCGGTAGTTG AAGAGGATCT AAAGCTTTTA   2280
      CAGCTTAAGC TAAGAGCCTC CGTGTCCACT AAATGTAACC TGGAAGACCA GGTAAAGAAA   2340
      TTGGAAGATG ACCGCAACTC ACTACAAGCT GCCAAAGCTG GACTGGAAGA TGAATGCAAA   2400
      ACCTTGAGGC AGAAAGTGGA GATTCTGAAT GAGCTCTATC AGCAGAAGGA GATGGCTTTG   2460
      CAAAAGAAAC TGAGTCAAGA AGAGTATGAA CGGCAAGAAA GAGAGCACAG GCTGTCAGCT   2520
      GCAGTAGAAA AGGCAGTTTC GGCTGCAGAG GAAGTAAAAA CTTACAAGCG GAGAATTGAA   2580
      GAAATGGAGG ATGAATTACA GAAGACAGAG CGGTCATTTA AAAACCAGAT CGCTACCCAT   2640
      GAGAAGAAAG CTCATGAAAA CTGGCTCAAA GCTCGTGCTG CAGAAAGAGC TATAGCTGAA   2700
      GAGAAAAGGG AAGCTGCCAA TTTGAGACAC AAATTATTAG AATTAACACA AAAGATGGCA   2760
      ATGCTGCAAG AAGAACCTGT GATTGTAAAA CCAATGCCAG AAAAACCAAA TACACAAAAC   2820
      CCTCCACGGA GAGGTCCTCT GAGCCAGAAT GGCTCTTTTG GCCCATCCCC TGTGAGTGGT   2880
      GGAGAATGCT CCCCTCCATT GACAGTGGAG CCACCCGTGA GACCTCTCTC TGCTACTCTC   2940
      AATCGAAGAG ATATGCCTAG AAGTGAATTT GGATCAGTGG ACGGGCCTCT ACCTCATCCT   3000
      CGATGGTCAG CTGAGGCATC TGGGAAACCC TCTCCTTCTG ATCCAGGATC TGGTACAGCT   3060
      ACCATGATGA ACAGCAGCTC AAGAGGCTCT TCCCCTACCA GGGTACTCGA TGAAGGCAAG   3120
      GTTAATATGG CTCCAAAAGG GCCCCCTCCT TTCCCAGGAG TCCCTCTCAT GAGCACCCCC   3180
      ATGGGAGGCC CTGTACCACC ACCCATTCGA TATGGACCAC CACCTCAGCT CTGCGGACCT   3240
      TTTGGGCCTC GGCCACTTCC TCCACCCTTT GGCCCTGGTA TGCGTCCACC ACTAGGCTTA   3300
```

320

```
AGAGAATTTG CACCAGGCGT TCCACCAGGA AGACGGGACC TGCCTCTCCA CCCTCGGGGA   3360
TTTTTACCTG GACACGCACC ATTTAGACCT TTAGGTTCAC TTGGCCCAAG AGAGTACTTT   3420
ATTCCTGGTA CCCGATTACC ACCCCCAACC CATGGTCCCC AGGAATACCC ACCACCACCT   3480
GCTGTAAGAG ACTTACTGCC GTCAGGCTCT AGAGATGAGC CTCCACCTGC CTCTCAGAGC   3540
ACTAGCCAGG ACTGTTCACA GGCTTTAAAA CAGAGCCCAT AAAACTATGA CCTCTGAGGT   3600
TTCATTGGAA AGAAAGTGTA CTGTGCATTA TCCATTACAG TAAAGGATTT CATTGGCTTC   3660
AAAATCCAAA AGTTTATTTT AAAAGGTTTG TTGTTAGAAC TAAGCTGCCT TGGCAGTGTG   3720
CATTTTTGAG CCAAACAATT CAAAAATGTC ATTTCTTCCC TAAATAAAAA TCACCTTTTA   3780
AGCTAGAGCG TCCTTACAAC TTTGAAATGT GCAATAAAGA ATACCTGTGT TTTAGCTAAT   3840
GTAGCATATG TAATTGCAAA ATGATTTAGA ATGTCATGAA AAATATGAAC ATTTCCTGTG   3900
GAAATGCTTT AAGAACATGT ATTTCCATTA TCCTATTTTT AGTGTACACC AGCTGAATAC   3960
GGAGCAATGG TGTTTATAAG CGTTTTTTTA AACTATCTGG TCACAAAGAC TGTTACGCTA   4020
AAAATGTTTA CTAAAAGATC ACTAAACTAT CTCCCCTCTT GCTGAAGTTC TTTGTAGTAA   4080
TAGCTCATAA AAATTTGTTT ATTAATATTT CCCAAGTGTC TGTTGACTCA TTGGACTGTT   4140
ATGAGGCTTG TGCCATTTGG GGAACATGTA AACTCAGGCT CCCAGAACTG AAGATGGTGG   4200
CTGGTGGCAC ACTTCCGGCT GCTCCTCCGT CACCTGTGAA CTCTACAAGT GATGTCTTTT   4260
TATTTCAAAG AAGTTTATTT CCCACTTGTA TAGCATTCAC ATGCTTTCTT TACGATCCTC   4320
ATTGTCTATT TGAGAATGGT TTTCTGAGAG TGAGTTTACA TTAGTAGCAA GAGTTGTTTG   4380
ACCTGATGTT CCATTGTTTT TACCATTCCT GTAGAAAAAG GGTGCACAAC AGAAAAATGA   4440
AAATGATGTG TCATGGCCAT AAAAGTATAG AAATCTTTAA AAATTTTAAA ATGTACAGTC   4500
CCTTATCTAT CTTTCCCATT CCTTGCCACT GATTTTTGAG GAATATAATA AAAAGATTGG   4560
AAGAGTATAA TGCCATGAGA AAGAATGATT TAGGACTGTG AGGGTTATAA CATGCCCTAG   4620
GTCAGCAACC AAGGGTTGAA ATCAGTTCTG TTTTAGGGGG AAATGGGGGG GGCGACAGAT   4680
ATTATTCCAA AATTAATATT AATTAATATT TAAACGTTGG TGTTTTTATT TAAAAATCAG   4740
TAACTAACCA TCTGGAATTG CACCATACTT AAAGTCTTAT CCATTACTAC ACTGTCTTTA   4800
AAACAATGTT TCTTTAAATA CTCTACAACG TTTCTAAGAA CGAACTTCAG ACATTTTAAT   4860
TACAGTAATA ATAGCACTCC TTTTAAGGAG TTTCAGATCC ACACTAAAAC TAAAATCATA   4920
AAAGGCTGAT ACTTTTGTTT GCTGCTAGGC TATATTCTTC CATTCTTTGA AGTCCTATGA   4980
TGTAATATTT TTGAAACCTA GTGTATGTCT TGTCACTGTT GTGATATTTA ATCGATTAAG   5040
AATACCTTGT AAAAAGGAGC AAAAGCTTCA ATGTGAAACA ATTTTCTCTC TTTATACTAA   5100
ACAACTGAAG ATAGATAGTT TAGAAAGATA AGGACCTTTG AAAGAAGACA ACTCTGTCAA   5160
AGTTCATAAG GAATATAAAA ATTCTTCAGG AAAAGAGAAT TCAATCTATA TGTCCTCCCG   5220
TTTAATATCA AGAATAGAAG AAATTAAGAG GAAAACTCCA CAGAAGAGCA TAGGCCACTT   5280
TTAGCCATGT AAAAATAAGA TTAAGTCACA AATACAACTT TTGAATTTAC CTGTCAATAT   5340
CTCTTTAGGA CACAAAACAA TGCTGAAGTT AATATAATTT CTAATTTTAA ATGTCATTTA   5400
AGTGTAGATT ATGCCATCTA GGAAGGTAAG TAGGAAAGGT AAATTAAATC TATTTTTAAA   5460
ATTCAAAATA TTAGAGTATT TTTCCCCTCT AAAGCCTTTT TTGGTGATTA TTCTGTATCT   5520
GACATAATTG AGAAACTGGT AAGCTGTAAA GATTCCAGTG TAGCTTCTCT GAGAAGTTGT   5580
GAGCCAGTCC ATAACTGCTT CCTCACATCC ATCTGATTGC ACCATTTCTG CAGCAAACCC   5640
CAAAGCAGGG TGCCAATATG CAGATGGCAT AGGGAGTATC ATCCCTCAGC CAAATCACTT   5700
TTCCATCTCT AAAGTTTCAT CTATTTTGGA AGTCATCTCC AACTAATTGT GTCTGGATTT   5760
AGTTGCTAAA ATTGTCTTAT TTATTTATGA AGCAGCAATA TTCAGCCTGA AAGCATTTCT   5820
GCCATAGTTG TTGTAGTTAT ATCGCCAATG GCTGATTTTT TTCATTGGAA AGTAAATTTA   5880
AGTAATTCGT GGGATGTGGT ATATTCTGTG TCAACTTCAA GATAATCACT CATTTTCTCG   5940
TTATATTCAG GTCTGAATTA AAGTTAAGTT AATCAC
```

SEQ ID NO:50 PAB7 Protein sequence
Protein Accession #:   BAA13448

```
1          11         21         31         41         51
|          |          |          |          |          |
AFLSKVEEDD YPSEELLEDE NAINAKRSKE KNPGNQGRQF DVNLQVPDRA VLGTIHPDPE    60
IEESKQETSM ILDSEKTSET AAKGVNTGGR EPNTMVEKER PLADKKAQRP FERSDFSDSI   120
KIQTPELGEV FQNKDSDYLK NDNPEEHLKT SGLAGEPEGE LSKEDHGNTE KYMGTESQGS   180
AAAEPEDDSF HWTPHTSVEP GHSDKREDLL IISSFFKEQQ SLQRFQKYFN VHELEALLQE   240
MSSKLKSAQQ ESLPYNMEKV LDKVFRASES QILSIAEKML DTRVAENRDL GMNENNIFEE   300
AAVLDDIQDL IYFVRYKHST AEETATLVMA PPLEEGLGGA MEEMQPLHED NFSREKTAEL   360
NVQVPEEPTH LDQRVIGDTH ASEVSQKPNT EKDLDPGPVT TEDTPMDAID ANKQPETAAE   420
EPASVTPLEN AILLIYSFMF YLTKSLVATL PDDVQPGPDF YGLPWKPVFI TAFLGIASFA   480
IFLWRTVLVV KDRVYQVTEQ QISEKLKTIM KENTELVQKL SNYEQKIKES KKHVQETRKQ   540
NMILSDEAIK YKDKIKTLEK NQEILDDTAK NLRVMLESER EQNVKNQDLI SENKKSIEKL   600
KDVISMNASE FSEVQIALNE AKLSEEKVKS ECHRVQEENA RLKKKKEQLQ QEIEDWSKLH   660
AELSEQIKSF EKSQKDLEVA LTHKDDNINA LTNCITQLNL LECESESEGQ NKGGNDSDEL   720
ANGEVGGDRN EKMKNQIKQM MDVSRTQTAI SVVEEDLKLL QLKLRASVST KCNLEDQVKK   780
LEDDRNSLQA AKAGLEDECK TLRQKVEILN ELYQQKEMAL QKKLSQEEYE RQEREHRLSA   840
ADEKAVSAAE EVKTYKRRIE EMEDELQKTE RSFKNQIATH EKKAHENWLK ARAAERAIAE   900
EKREAANLRH KLLELTQKMA MLQEEPVIVK PMPGKPNTQN PPRRGPLSQN GSFGPSPVSG   960
GECSPPLTVE PPVRPLSATL NRRDMPRSEF GSVDGPLPHP RWSAEASGKP SPSDPGSGTA  1020
TMMNSSSRGS SPTRVLDEGK VNMAPKGPPP FPGVPLMSTP MGGPVPPPIR YGPPPQLCGP  1080
FGPRPLPPPF GPGMRPPLGL REFAPGVPPG RRDLPLHPRG FLPGHAPFRP LGSLGPREYF  1140
IPGTRLPPPT HGPQEYPPPP AVRDLLPSGS RDEPPPASQS TSQDCSQALK QSP
```

SEQ ID NO:51 PAB9 DNA SEQUENCE

Nucleic Acid Accession #:   NM_006457
Coding sequence: 84-1874 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
AGACTGAGGC GGAGGCAGCC CCGCGCCGCG CCGGACCCGA GCATATTTCA TTTTCTGTCA    60
```

321

```
TTGGACTTTG AGCCATTAGA ACCATGAGCA ACTACAGTGT GTCACTGGTT GGCCCAGCTC   120
CTTGGGGTTT CCGGCTGCAG GGCGGTAAGG ATTTCAACAT GCCTCTGACA ATCTCTAGTC   180
TAAAAGATGG CGGCAAGGCA GCCCAGGCAA ATGTAAGAAT AGGCGATGTG GTTCTCAGCA   240
TTGATGGAAT AAATGCACAA GGAATGACTC ATCTTGAAGC CCAGAATAAG ATTAAGGGTT   300
GTACAGGCTC TTTGAATATG ACTCTGCAAA GAGCATCTGC TGCACCCAAG CCTGAGCCGG   360
TTCCTGTTCA AAAGGGAGAA CCTAAAGAAG TAGTTAAACC TGTGCCCATT ACATCTCCTG   420
CTGTGTCCAA AGTCACTTCC ACAAACAACA TGGCCTACAA TAAGGCACCA CGGCCTTTTG   480
GTTCTGTGTC TTCACCAAAA GTCACATCCA TCCCATCACC ATCGTCTGCC TTCACCCCAG   540
CCCATGCGAC CACCTCATCA CATGCTTCCC CTTCACCCGT GGCTGCCGTC ACTCCTCCCC   600
TGTTCGCTGC ATCTGGACTG CATGCTAATG CCAATCTTAG TGCTGACCAG TCTCCATCTG   660
CACTGAGCGC TGGTAAAACT GCAGTTAATG TCCCACGGCA GCCCACAGTC ACCAGCGTGT   720
GTTCCGAGAC TTCTCAGGAG CTAGCAGAGG GACAGAGAAG AGGATCCCAG GGTGACAGTA   780
AACAGCAAAA TGGCCCACCA AGAAACACA TTGTGGAGCG CTATACAGAG TTTTATCATG   840
TACCCACTCA CAGTGATGCC AGCAAGAAGA GACTGATTGA GGATACTGAA GACTGGCGTC   900
CAAGAACTGG AACAACTCAG TCTCGCTCTT TCCGAATCCT TGCCCAGATC ACTGGGACTG   960
AACATTTGAA AGAATCTGAA GCCGATAATA CAAAGAAGGC AAATAACTCT CAGGAGCCTT  1020
CTCCGCAGTT GGCTTCCTTG GTAGCTTCCA CACGGAGCAT GCCCGAGAGC CTGGACAGCC  1080
CAACCTCTGG CAGACCAGGG GTTACCAGCC TCACAACTGC AGCTGCCTTC AAGCCTGTAG  1140
GATCCACTGG CGTCATCAAG TCACCAAGCT GGCAACGGCC AAACCAAGGA GTACCTTCCA  1200
CTGGAAGAAT CTCAAACAGC GCTACTTACT CAGGATCAGT GGCACCAGCC AACTCAGCTT  1260
TGGGACAAAC CCAGCCAAGT GACCAGGACA CTTTAGTGCA AAGAGCTGAG CACATTCCAG  1320
CAGGGAAACG AACTCCGATG TGCGCCCATT GTAACCAGGT CATCAGAGGA CCATTCTTAG  1380
TGGCACTGGG GAAATCTTGG CACCCAGAAG AATTCAACTG CGCTCACTGC AAAAATACAA  1440
TGGCCTACAT TGGATTTGTA GAGGAGAAAG GAGCCCTGTA TTGTGAGCTG TGCTATGAGA  1500
AATTCTTTGC CCCTGAATGT GGTCGATGCC AAAGGAAGAT CCTTGGAGAA GTCATCAATG  1560
CGTTGAAACA AACTTGGCAT GTTTCCTGTT TTGTGTGTGT AGCCTGTGGA AAGCCCATTC  1620
GGAACAATGT TTTTCACTTG GAGGATGGTG AACCCTACTG TGAGACTGAT TATTATGCCC  1680
TCTTTGGTAC TATATGCCAT GGATGTGAAT TTCCCATAGA AGCTGGTGAC ATGTTCCTGG  1740
AAGCTCTGGG CTACACCTGG CATGACACTT GCTTTGTATG CTCAGTGTGT GTGAAAGTT   1800
TGGAAGGTCA GACCTTTTTC TCCAAGAAGG ACAAGCCGCT GTGTAAGAAA CATGCTCATT  1860
CTGTGAATTT TTGAAAGTCA ACAGTTCAGG AGAAGAGAAG GAATTTGAAG AGAAAAAGGA  1920
AAATTAAAAT TACTAATTAA TTTTTAGATT CAATATTTAT ATGGAGTTTT GAAAAATAAT  1980
AGTGGCCCTG AAGGAATAAA TTCCAGCTTT AAAAACCAAG TCTGAGGAAA TATTTGGCTT  2040
CATAAAGTAA AGAGACGGTT TGGCATTTAT TATTACTTTT TCCTGTATTT TATGCCCATA  2100
AAATAAGCTT TATAAAAACC AATTTCCTGA TGGACTATTA AATTCATCTT AGAATAAATT  2160
AGTGAAGAAT TTAATTTTAG AATAAATAAT CCAATCTGAA ATAATTATAC CTTCTTTCCT  2220
TGTTAGGTAG TTATGAGTAA ATCTGCAAAA GGCAATGAAA ATGCCTTAAA TTTTATCAAT  2280
AACAGAATTA TTGTATTTAA AAAAAAACTA ATACTTATCT TTAAAATAGT AAATAGGATT  2340
TTAAACAGAG AATTTTATCA GTAATAGGTG TCAGTTTTTA AAAAATTGCT TGTAGGCTGA  2400
GCGCGGTGGC TCACGCCTGT AATCCCAGCA CTTTGGGAGG CCAAGGTGGG TGGACCACAT  2460
GAGGTCAGGA GTTTGAGATC AGCCTGGCCA ACATGGTGAA ACCCCATCTC TACTAAAAAT  2520
ACAAAAATTA GCCGGACGCA GTGGCACGCG CCTGTAATCC CAGCTACTCA AGAGGCTGAG  2580
GCACGAGAAT CACTTGAACC CGGGAGGGGA AGGTTGCAGT GAGCCAAGAT CGTACCACTG  2640
CACTCCAGCC TGGGTGACAG AGTGAGACTC CGTCTCCAAA AAAAAACTTT GCTTGTATAT  2700
TATTTTTGCC TTACAGTGGA TCATTCTAGT AGGAAAGGAC AATAAGATTT TTTATCAAAA  2760
TGTGTCATGC CAGTAAGAGA TGTTATATTC TTTTCTTATT TCTTCCCCAC CCAAAAATAA  2820
GCTACCATAT AGCTTATAAG TCTCAAATTT TTGCCTTTTA CTAAAATGTG ATTGTTTCTG  2880
TTCATTGTGT ATGCTTCATC ACCTATATTA GGCAAATTCC ATTTTTTCCC TTGCGCTAAG  2940
GTAAAGATTT AATTAAATAA TTTTTGGCCTC TCATAGTTTT CTCTCTCTTT AAAGAGAATA  3000
AATAGAGGGC CAGGTGTGGT GGCTCACGCC TGTGATCCCA GCACTTTGGG AGGCCAAGAC  3060
GGGCGGATCA TGAGGTCAAG AGATCAAGAT CATCCTGGCC AACATGGTGA AACCCTGTCT  3120
CTACTAAAAA TACAAAAATG AGCTGGGCAT GGTGGGGCGT GCCTGTAGTC CCATGTACTT  3180
GGGAGGCTGA GGCAGGAAAA TTCTTGAACC CAGGAGACGG AAGTTGCAGT GAGCTGAGAT  3240
CACACCACTG CACTCCAGCC TGGTGACAGA GCAAGACTCC GGCTCTT
```

SEQ ID NO:52 PAB9 Protein sequence
Protein Accession #:   NP_006448

```
       1         11        21        31        41        51
       |          |         |         |         |         |
   1 MSNYSVSLVG PAPWGFRLQG GKDFNMPLTI SSLKDGGKAA QANVRIGDVV LSIDGINAQG   60
  61 MTHLEAQNKI KGCTGSLNMT LQRASAAPKP EPVPVQKGEP KEVVKPVPIT SPAVSKVTST  120
 121 NNMAYNKAPR PFGSVSSPKV TSIPSPSSAF TPAHATTSSH ASPSPVAAVT PPLFAASGLH  180
 181 ANANLSADQS PSALSAGKTA VNVPRQPTVT SVCSETSQEL AEGQRRGSQG DSKQQNGPPR  240
 241 KHIVERYTEF YHVPTHSDAS KKRLIEDTED WRPRTGTTQS RSFRILAQIT GTEHLKESEA  300
 301 DNTKKANNSQ EPSPQLASLV ASTRSMPESL DSPTSGRPGV TSLTTAAAFK PVGSTGVIKS  360
 361 PSWQRPNQGV PSTGRISNSA TYSGSVAPAN SALGQTQPSD QDTLVQRAEH IPAGKRTPMC  420
 421 AHCNQVIRGP FLVALGKSWH PEEFNCAHCK NTMAYIGFVE EKGALYCELC YEKFFAPECG  480
 481 RCQRKILGEV INALKQTWHV SCFVCVACGK PIRNNVFHLE DGEPYCETDY YALFGTICHG  540
 541 CEFPIEAGDM FLEALGYTWH DTCFVCSVCC ESLEGQTFFS KKDKPLCKKH AHSVNF
```

SEQ ID NO:53 PBH7 DNA SEQUENCE

Nucleic Acid Accession #:   AA431407
Coding sequence: 1-864 (underlined sequences correspond to start and stop codons)

```
1         11        21        31        41        51
|          |         |         |         |         |
ATGGCCAACT GTAAAATGAC CAAAAGCATC AGGTTCCCTG CCCTGGAGCA CTGCTATACT    60
GGCGGGGAGG TCGTGTTGCC CAAGGATCAG GAGGAGTGGA AAGACGGAC GGGCCTTCTG   120
CTCTACGAGA ACTATGGGCA GTCGGAAACG GGACTAATTT GTGCCACCTA CTGGGGAATG   180
```

```
AAGATCAAGC CGGGTTTCAT GGGGAAGGCC ACTCCACCCT ATGACGTCCA GTTTCATATG    240
GAGGCCTCAG TTGAAAACTG CATTATTGTG AGCATGAACA CCGCTGACCC TGGCAGCCAG    300
GGCATCACAC ACAGCCTCTT GCTACAGGTC ATTGATGACA AGGGCAGCAT CCTGCCACCT    360
AACACAGAAG GAAACATTGG CATCAGAATC AAACCTGTCA GGCCTGTGAG CCTCTTCATG    420
TGCTATGAGG GTGACCCAGA GAAGACAGCT AAAGTGGAAT GTGGGGACTT CTACAACACT    480
GGGGACAGAG GAAAGATGGA TGAAGAGGGC TACATTTGTT TCCTGGGGAG GAGTGATGAC    540
ATCATTAATG CCTCTGGGTA TCGCATCGGG CCTGCAGAGG TTGAAAGCGC TTTGGTGGAG    600
CACCCAGCGG TGGCGGAGTC AGCCGTGGTG GGCAGCCCAG ACCCGATTCG AGGGGAGGTG    660
GTGAAGGCCT TTATTGTCCT GACCCCACAG TTCCTGTCCC ATGACAAGGA TCAGCTGACC    720
AAGGAACTGC AGCAGCATGT CAAGTCAGTG ACAGCCCCAT ACAAGTACCC AAGGAAGGTG    780
GAGTTTGTCT CAGAGCTGCC AAAAACCATC ACTGGCAAGA TTGAACGGAA GGAACTTCGG    840
AAAAAGGAGA CTGGTCAGAT GTAATCGGCA GTGAACTCAG AACGCACTGC ACACCTGAGG    900
CAAATCCCTG GCCACTTTAG TCTCCCCACT ATGGTGAGGA CGAGGGTGGG GCATTGAGAG    960
TGTTGATTTG GGAAAGTATC AGGAGTGCCA TGATTCCAAT GTTTTCCTTC TTTTAAATTA   1020
AATTCAGTTG CTCTGCTTCC TCCAAGTCCT CTGTATCTTT AGAATTTCCC AGGTGAGCAC   1080
TCATAACGCA AGTAATAAAA TACTGATATC AACAA
```

SEQ ID NO:54 PBH7 Protein sequence
Protein Accession #:   FGENESH predicted

```
1          11         21         31         41         51
|          |          |          |          |          |
MANCKMTKSI RFPALEHCYT GGEVVLPKDQ EEWKRRTGLL LYENYGQSET GLICATYWGM    60
KIKPGFMGKA TPPYDVQFHM EASVENCIIV SMNTADPGSQ GITHSLLLQV IDDKGSILPP   120
NTEGNIGIRI KPVRPVSLFM CYEGDPEKTA KVECGDFYNT GDRGKMDEEG YICFLGRSDD   180
IINASGYRIG PAEVESALVE HPAVAESAVV GSPDPIRGEV VKAFIVLTPQ FLSHDKDQLT   240
KELQQHVKSV TAPYKYPRKV EFVSELPKTI TGKIERKELR KKETGQM
```

SEQ ID NO:55 PBJ5 DNA SEQUENCE
Nucleic Acid Accession #:   AF388200
Coding sequence: 33-137 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
GAGAGAGGGA GGCAGAAGAG GAAGTCAGAG CGATGTGCTG TGAAATCTAC TACCGTTTGC    60
TGGTTTTGAA AATGGAGAAA AAGAGTGAGG AACTGAGAAA CATGGATGGC CTTGGGAACG   120
TGGAAAGGG TCACTGAAAT GGGACGACAT GAACTCAAGG AGGCTATTTA TGACCATGTC   180
ATTTGCAACA TGAAGAAAGC TTATCTGGAG TGAAAGTAAA TGAGACCAAC AGAGATAAGA   240
GACCCGGAGA AATCCTGGTT ACACTGCTTG AATCCTGTCA GTCCTATACT GGAGTCCTGT   300
TAATACAAAA TAATAGTAAT AATCCCTCTG TTTCTTATGT TTATGCCAAC TTCAACAAAA   360
AGAAACTTGA CTAAGAGACA ATATAAGAAC TTAATGTGTA ATTAAGAAAG AACTCTCCAC   420
CACGGGGAAT GTGAAGGTA TATGAGTCCC TTTTCACGAT GCGATGTCAT GTCTTTTAAA   480
TAAGCCATAC TTTATGTTCA ATAAAAAGAG AATAAGCAGG A
```

SEQ ID NO:56 PBJ5 Protein sequence
Protein Accession #:   AAK83352

```
1          11         21         31         41         51
|          |          |          |          |          |
MCCEIYYRLL VLKMEKKSEE LRNMDGLGNV EKGH
```

SEQ ID NO:57 PBJ7 DNA SEQUENCE
Nucleic Acid Accession #:   AA876910
Coding sequence: 1-2064 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGGACAGTT GCCTGCAACA TATGAGAGAC CTACTTTACC TCCTTCAGGA GCTCAGGTGT    60
TTAAATCCAG CTACACTACT CCCTGATCCA GACTCCACTA CTCCTGTTCA TGACTGTCAG   120
GATCTGTTGG AAACTACCAA AACTGGCCAA CCTGATCTTC AAGATGTGCC CCTAGAAAAG   180
GCAGATGCCA CTGTGTTCAC AGATGGTAGC AGCTTCCTCG AGCAGGGAGA ACGAAAAGCT   240
GTTTCTTTTC CACAGCCAGA TCTGCCTGAC AATCCCACAT ACTCAACAGA AGAAGAAAAA   300
CTGGCTTCAG ATGTTGGAGC AAATAAAAAT CAGGAAGGAC GTGTATTCGC AAACACTACT   360
TGGAGGGCCG GTACCTCCAA GGAAGTCTCC TTTGCAGTTG ATTTATGTGT ACTGTTCCCA   420
GAGCCAGCTC GTACCCATGA AGAGCAACAT AATTTGCCGG TCATAGGAGC AGGAAGTGTC   480
GACCTTGCAG CAGGATTTGG ACACTCTGGG AGCCAAACTG GATGTGGAAG CTCCAAAGGT   540
GCAGAAAAAG GGCTCCAAAA TGTTGACTTT TACCTCTGTC CTGGAAATCA CCCTGACGCT   600
AGCTGTAGAG ATACTTACCA GTTTTTCTGC CCTGATTGGA CATGTGTAAC TTTAGCCACC   660
TACTCTGGGG GATCAACTAG ATCTTCAACT CTTTCCATAA GTCGTGTTCC TCATCCTAAA   720
TTATGTACTA GAAAAAATTG TAATCCTCTT ACTATAACTG TCCATGACCC TAATGCAGCT   780
CAATGGTATT ATGGCATGTC ATGGGGATTA AGACTTTATA TCCCAGGATT TGATGTTGGG   840
ACTATGTTCA CCATCCAAAA GAAAATCTTG GTCTCATGGA GCTCCCCCAA GCCAATCGGG   900
CCTTTAACTG ATCTAGGTGA CCCTATATTC CAGAAACACC CTGACAAAGT TGATTTAACT   960
GTTCCTCTGC CATTCTTAGT TCCTAGACCC CAGCTACAAC AACAACATCT TCAACCCAGC   1020
CTAATGTCTA TACTAGGTGG AGTACACCAT CTCCTTAACC TCACCCAGCC TAAACTAGCC   1080
CAAGATTGTT GGCTATGTTT AAAAGCAAAA CCCCCTTATT ATGTAGGATT AGGAGTAGAA   1140
GCCACACTTA AACGTGGCCC TCTATCTTGT CATACACGAC CCCGTGCTCT CACAATAGGA   1200
GATGTGTCTG GAAATGCTTC CTGTCTGATT AGTACCGGGT ATAACTTATC TGCTTCTCCT   1260
TTTCAGGCTA CTTGTAATCA GTCCCTGCTT ACTTCCATAA GCACCTCAGT CTCTTACCAA   1320
GCACCCAACA ATACCTGGTT GGCCTGCACC TCAGGTCTCA CTCGCTCCAA TAATGGAACT   1380
```

```
GAACCAGGAC CTCTCCTGTG CGTGTTAGTT CATGTACTTC CCCAGGTATA TGTGTACAGT  1440
GGACCAGAAG GACGACAACT CATCGCTCCC CCTGAGTTAC ATCCCAGGTT GCACCAAGCT  1500
GTCCCACTTC TGGTTCCCCT ATTGGCTGGT CTTAGCATAG CTGGATCAGC AGCCATTGGT  1560
ACGGCTGCCC TGGTTCAAGG AGAAACTGGA CTAATATCCC TGTCTCAACA GGTGGATGCT  1620
GATTTTAGTA ACCTCCAGTC TGCCATAGAT ATACTACATT CCCAGGTAGA GTCTCTGGCT  1680
GAAGTAGTTC TTCAAAACTG CCGATGCTTA GATCTGCTAT TCCTCTCTCA AGGAGGTTTA  1740
TGTGCAGCTC TAGGAGAAAG TTGTTGCTTC TATGCCAATC AATCTGGAGT CATAAAAGGT  1800
ACAGTAAAAA AAGTTCGAGA AAATCTAGAT AGGCACCAAC AAGAACGAGA AAATAACATC  1860
CCCTGGTATC AAAGCATGTT TAACTGGAAC CCATGGCTAA CTACTTTAAT CACTGGGTTA  1920
GCTGGACCTC TCCTCATCCT ACTATTAAGT TTAATTTTTG GGCCTTGTAT ATTAAATTCG  1980
TTTCTTAATT TTATAAAACA ACGCATAGCT TCTGTCAAAC TTACGTATCT TAAGACTCAA  2040
TATGACACCC TTGTTAATAA CTGA
```

```
1          11         21         31         41         51
|          |          |          |          |          |
MDSCLQHMRD LLYLLQELRC LNPATLLPDP DSTTPVHDCQ DLLETTKTGQ PDLQDVPLEK   60
ADATVFTDGS SFLEQGERKA VSFPQPDLPD NPTYSTEEEK LASDVGANKN QEGRVFANTT  120
WRAGTSKEVS FAVDLCVLFP EPARTHEEQH NLPVIGAGSV DLAAGFGHSG SQTGCGSSKG  180
AEKGLQNVDF YLCPGNHPDA SCRDTYQFFC PDWTCVTLAT YSGGSTRSST LSISRVPHPK  240
LCTRKNCNPL TITVHDPNAA QWYYGMSWGL RLYIPGFDVG TMFTIQKKIL VSWSSPKPIG  300
PLTDLGDPIF QKHPDKVDLT VPLPFLVPRP QLQQQHLQPS LMSILGGVHH LLNLTQPKLA  360
QDCWLCLKAK PPYYVGLGVE ATLKRGPLSC HTRPRALTIG DVSGNASCLI STGYNLSASP  420
FQATCNQSLL TSISTSVSYQ APNNTWLACT SGLTRCINGT EPGPLLCVLV HVLPQVYVYS  480
GPEGRQLIAP PELHPRLHQA VPLLVPLLAG LSIAGSAAIG TAALVQGETG LISLSQQVDA  540
DFSNLQSAID ILHSQVESLA EVVLQNCRCL DLLFLSQGGL CAALGESCCF YANQSGVIKG  600
TVKKVRENLD RHQQERENNI PWYQSMFNWN PWLTTLITGL AGPLLILLLS LIFGPCILNS  660
FLNFIKQRIA SVKLTYLKTQ YDTLVNN
```

```
1          11         21         31         41         51
|          |          |          |          |          |
TGATGGTGGA AATTTCTTGA AACCGCTCTC GTAATTTGCC ACGTGCTGTT GCAAATATTC   60
TGGTGAATGA ACACAGAATC AGCATGGCTT TCCTTTGCTG AGAAATCACT GATGGGAAGT  120
GAGACTTGTT AAACTTGAAA GTGAATGGAC CTGAGTGGAC CCTTTGATCA CATCAGTAAA  180
CATGAGCGGT ACCAAACCTG ATATTTTATG GGCACCACAC CATGTTGATA GATTTGTTGT  240
GTGTGACTCA GAACTAAGTC TTTATCATGT GGAATCTACT GTGAATTCAG AACTCAAAGC  300
TGGATCTTTA CGTTTATCTG AAGACTCTGC AGCTACATTA CTGTCAATAA ATTCAGATAC  360
ACCCTATATG AAATGTGTTG CCTGGTATCT TAATTATGAT CCTGAATGTC TGCTGGCAGT  420
TGGACAAGCA AATGGTCGAG TTGTACTTAC AAGCCTTGGT CAAGATCATA ACTCAAAGTT  480
CAAAGATTTG ATAGGAAAAG AGTTTGTTCC AAAACATGCA CGACAATGTA ATACCCTTGC  540
CTGGAATCCA CTGGATAGTA ACTGGCTAGC TGCTGGTTTA GATAAGCACA GAGCTGACTT  600
TTCAGTGCTA ATATGGGATA TCTGCAGCAA ATATACTCCT GATATAGTTC CCATGGAAAA  660
AGTGAAACTT TCAGCAGGTG AAACTGAAAC AACATTATTA GTAACAAAAC CACTTTATGA  720
GTTAGGACAG AATGATGCTT GTCTGTCTCT TTGTTGGCTT CCACGAGACC AGAAACTTCT  780
CCTTGCTGGT ATGCATCGTA ACCTAGCTAT ATTTGATCTT CGGAATACAA GCCAAAAGAT  840
GTTCGTAAAT ACAAAAGCTG TTCAGGGTGT GACGGTAGAC CCATATTTCC ACGATCGTGT  900
TGCTTCCTTC TATGAAGGTC AGGTTGCAAT ATGGGATCTT AGAAAATTTG AGAAGCCAGT  960
TTTGACATTG ACTGAGCAAC CAAAACCCTT AACAAAAGTA GCATGGTGTC CCACTAGGAC 1020
TGGTCTACTT GCCACTTTAA CAAGGGATAG TAATATTATT AGATTGTATG ATATGCAGCA 1080
TACACCCACT CCCATTGGGG ATGAAACTGA ACCCACAATA ATTGAAAGAA GTGTGCAACC 1140
TTGTGACAAT TACATTGCTT CCTTTGCGTG GCATCCAACA AGTCAAAATC GAATGATAGT 1200
TGTAACTCCC AACCGAACAA TGTCAGACTT CACTGTTTTT GAAAGGATAT CTCTTGCCTG 1260
GAGCCCAATT ACATCTTTAA TGTGGGCTTG TGGTCGTCAT TTATATGAAT GTACGGAAGA 1320
AGAAAATGAT AATTCTTTAG AAAAAGATAT AGCAACGAAG ATGCGTCTTC GGGCTTTATC 1380
AAGGTATGGA CTTGATACAG AGCAGGTGTG GAGGAACCAC ATTTTAGCTG GAAATGAAGA 1440
TCCACAGCTC AAGTCACTCT GGTATACTCT GCACTTTATG AAGCAATACA CAGAAGATAT 1500
GGATCAGAAA TCTCCAGGCA ACAAAGGATC ATTGGTTTAT GCAGGAATTA AATCAATTGT 1560
AAAGTCATCG TTGGGAATGG TGGAAAGCAG CAGACATAAT TGGAGTGGGT TGGATAAGCA 1620
AAGTGATATT CAAAACTTAA ATGAAGAGAG AATCTTAGCT TTACAGCTTT GTGGGTGGAT 1680
AAAGAAAGGA ACGGATGTAG ACGTGGGGCC ATTTTTGAAC TCCCTTGTAC AAGAAGGGGA 1740
ATGGGAAAGA GCTGCTGCTG TGGCATTGTT CAACTTGGAT ATTCGCCGAG CAATCCAAAT 1800
CCTGAATGAA GGGGCATCTT CTGAAAAAGG CAGGAGATCT GAATCTCAAT GTGGTAGCAA 1860
TGGCTTTATC GGGTTATACG GATGAGAAGA ACTCCCTTTG GAGAGAAATG TGTAGCACAC 1920
TGCGATTACA GCTAAATAAC CCGTATTTGT GTGTCATGTT TGCATTTCTG ACAAGTGAAA 1980
CAGGATCTTA CGATGGAGTT TTGTATGAAA ACAAAGTTGC AATACGTGAC AGAGTGGCAT 2040
TTGCTTGTAA ATTCCTTAGT GATACTCAGA TACATCGAAA AGTTGACCAA TGAAATGAAA 2100
GAGGCTGGAA ATTTGGAAGG AATTTTGCTT ACAGGCCTTA CTAAAGATGG AGTGGACTTA 2160
ATGGAGAGTT ATGTTGATAG AACTGGAGAT GTTCAAACAG CAAGTTACTG TATGTTACAG 2220
GGTTCACCTT TAGATGTTCT TAAAGATGAA AGGGTTCAGT ACTGGATTGA GAATTATAGA 2280
AATTTATTAG ATGCCTGGAG GTTTTGGCAT AAACGAGCTG AATTTGTATT TCACAGGAGT 2340
AAGTTGGATC CCAGTTCCAA GCCTTTAGCA CAAGTTTTTG TGAGTTGCAA TTTCTGTGGC 2400
AAGTCAATCT CCTACAGCTG TTCAGCTGTG CCTCATCAGG GCAGAGGTTT TAGTCAGTAT 2460
GGTGTGAGTG GCTCACCAAC GAAATCTAAA GTCACAAGTT GTCCTGGCTG TCGAAAACCA 2520
CTTCCTCGAT GTGCGCTTTG TCTCATTAAT ATGGGAACAC CAGTTTCTAG CTGTCCTGGA 2580
```

```
        GGAACCAAAT CAGATGAAAA AGTGGACTTG AGCAAGGACA AAAAATTAGC CCAATTTAAC   2640
        AACTGGTTTA CATGGTGTCA TAATTGCAGG CACGGTGGAC ATGCTGGACA TATGCTTAGT   2700
        TGGTTCAGGG ACCATGCAGA GTGCCCTGTG TCTGCATGCA CGTGTAAATG TATGCAGTTG   2760
  5     GATACAACGG GGAATCTGGT ACCTGCAGAG ACTGTCCAGC CATAAAATGT TACCACCTTA   2820
        AGAGAACCCT TCAAGTGTGG AGCTTTCTAG TAGGTGTCCT TCATAGCTCA GAAACATACC   2880
        TCAGAACAAG CCATTCATGA CTTACCTGTA ATGGGAAAAT AAATCATTCT ATCAGAAAAA   2940
        AAAAAAAAAA AAAAAAAAAA
```

SEQ ID NO:60 PCQ1 Protein sequence
Protein Accession #:   NP_061878

```
        1          11         21         31         41         51
        |          |          |          |          |          |
        MSGTKPDILW APHHVDRFVV CDSELSLYHV ESTVNSELKA GSLRLSEDSA ATLLSINSDT   60
        PYMKCVAWYL NYDPECLLAV GQANGRVVLT SLGQDHNSKF KDLIGKEFVP KHARQCNTLA   120
        WNPLDSNWLA AGLDKHRADF SVLIWDICSK YTPDIVPMEK VKLSAGETET TLLVTKPLYE   180
        LGQNDACLSL CWLPRDQKLL LAGMHRNLAI FDLRNTSQKM FVNTKAVQGV TVDPYFHDRV   240
        ASFYEGQVAI WDLRKFEKPV LTLTEQPKPL TKVAWCPTRT GLLATLTRDS NIIRLYDMQH   300
        TPTPIGDETE PTIIERSVQP CDNYIASFAW HPTSQNRMIV VTPNRTMSDF TVFERISLAW   360
        SPITSLMWAC GRHLYECTEE ENDNSLEKDI ATKMRLRALS RYGLDTEQVW RNHILAGNED   420
        PQLKSLWYTL HFMKQYTEDM DQKSPGNKGS LVYAGIKSIV KSSLGMVESS RHNWSGLDKQ   480
        SDIQNLNEER ILALQLCGWI KKGTDVDVGP FLNSLVQEGE WERAAAVALF NLDIRRAIQI   540
        LNEGASSEKG RRSESQCGSN GFIGLYG
```

SEQ ID NO:61 PDG3 DNA SEQUENCE
Nucleic Acid Accession #:   U42359
Coding sequence: 563-775 (underlined sequences correspond to start and stop codons)

```
        1          11         21         31         41         51
        |          |          |          |          |          |
        TTGTACATCT TAACAACCTT AAGCTGTACA AATAGANCAA TAATATCTAA ATGGTGTGAT   60
        GATCAGCCCA CAGTACACAT CATTGATGAG AATTTCACTG GTCTCAACCT TTCTCATGCT   120
        GAGTCCTGGC TTTGTAAAAT GACTTATAAA GGTCCAAGGA TTTAGAGATG ATTAAGAGAT   180
        AAGCTGGCAT TCTGTAAAGG CACCATCGTC TATCCCCTGT CTTATCTAGA TAAAGAATGT   240
        AGTGCTAAAT CTTGTAATAA TATTGTACAA ATGGAAATTC AATCTTAAGG ATTATTTTTT   300
        CCATATTGTT GTATTTCATT GTGGTGTATT GGAAAGTGAT CTGGACTTTG AGTGAGAAGA   360
        TGTGATTTGG ACCATGGCAC TTAAAAACTC TATAACCTCA GGCAAGTCTT TTAATCTTCT   420
        CTGAGCCTCA GTTTTCCTCA TTTTTCAAAT ATAGAGAGTA TAACATTTAT CTCATAAGAC   480
        AAGTTGTAGT AAATTACTGT TTTACAAATG TAAGATAACT TTTAACTGTG AGATTCCATA   540
        TTCCAGTCTT ACATTATT̲A̲T̲ GTTTATCTGC CACAGGGAGA AGTCCTCAGA TAAAAATGTC   600
        TACCAAAAGA CTGACACGTG GAGTTAATCA TTTGACAGAT GCAAATGCTT CCACCCCCAA   660
        CAAATATACT TTCTTTAACT TCTGTGTGGG TATCACTTAG GGAAAAAAAG GCAGGCAACA   720
        AAATATTTTT TAATTCTATC TTAGGAAAAA TTGTAGNCAA ATCTTTTTNT CCCATT̲A̲A̲C̲A̲   780
        AATAATGTAA GCCTTAATAT TCAAGGGGTA ATAAAAATAC AAAGTCTTCC AAACAGGTAA   840
        CTTACTTGAA AACTTT
```

SEQ ID NO:62 PDG3 Protein sequence
Protein Accession #:   AAB18375

```
        1          11         21         31         41         51
        |          |          |          |          |          |
        MGARGAPSRR RQAGRRLRYL PTGSFPFLLL LLLLCIQLGG GQKKKENLLA EKVEQLMEWS   60
        SRRSIFRMNG DKFRKFIKAP PRNYSMIVMF TALQPQRQCS VCRQANEEYQ ILANSWRYSS   120
        AFCNKLFFSM VDYDEGTDVF QQLNMNSAPT FXHXPPKGRP KRADTFDLQR IGFAAEQLAK   180
        WIADRTDVHI RVFRPPNYSG TIALALLVSL VGGLLYXRRN NLEFIYNKTG WAMVSLCIVF   240
        AMTSGQMWNH IRGPPYAHKN PHNGQVSYIH GSSQAQFVAE SHIILVLNAA ITMGMVLLNE   300
        AATSKGDVGK RRIICLVGLG LVVFFFSFLL SIFRSKYHGY PYSDLDFE
```

SEQ ID NO:63 PDG8 DNA SEQUENCE
Nucleic Acid Accession #:   AL080235
Coding sequence: 245-453 (underlined sequences correspond to start and stop codons)

```
        1          11         21         31         41         51
        |          |          |          |          |          |
        GGTCGCCGCA CCGGCCGCCT CCGGCCCGCC GCCGCCCCCA GCGCCGCCGC CGCCACCGCC   60
        GGGGCGCCCA CCGCGCTGCC AGCCTACCCC GCGGCCGAGC CGCCCGGGCC GCTGTGGCTG   120
        CAGGGCGAGC CGCTGCATTT CTGCTGCCTA GACTTCAGCC TGGAGGAGCT GCAGGGCGAG   180
        CCGGGCTGGC GGCTGAACCG TAAGCCCATT GAGTCCACGC TGGTGGCCTG CTTC̲A̲T̲GACC   240
        CTGGTCATCG TGGTGTGGAG CGTGGCCGCC CTCATCTGGC CGGTGCCCAT CATCGCCGGC   300
        TTCCTGCCCA ACGGCATGGA ACAGCGCCGG ACCACCGCCA GCACCACCGC AGCCACCCCC   360
        GCCGCAGTGC CCGCAGGGAC CACCGCAGCC GCCGCCGCCG CCGCCGCTGC CGCCGCCGCC   420
        GCGGCCGTCA CTTCGGGGGT GGCGACCAAG T̲G̲A̲CCCGCTC CGCTCCTCCC TGTGTCCGTC   480
        CTGTGTCCGC GCGCGCGGGT GCCTTTCCCG CCGGGGACTC GGCCGGTGTG CTTCGTGCTG   540
        TAGTTATCGT TAGTTCCTCT TCCCGAGATG GGGCCGCCGA GAGACCCCAG CGCCTTTGAA   600
        AAGCAAGGTT TGTGCTGCGC TTCCAGTTCC GAAAAGCAGA TGTTTAAGCC CTTGGACTGA   660
        GGGTGGGATC GCAGCTCCGA AGACGGAGAG GAGGGAAATG GGGCCCTTTC CCCTCTATTG   720
        CATCCCCCTG CCCGACTCCT TCCCCGCACC CACGTGCCCT AGATTCATGG CAGAAAATGA   780
        CCAAATCCTG TGTATTTGTT TTATATATTT AATAACTGTT TTAAATGAAA GTTTTAGTAA   840
        AAAAAATACA AAACAAAAAG ATTAAATTGC TATTGCTGTA GTAAGAGAAG CTCTTTGTAT   900
        CTGAACATAG TTGTATTTGA AATTTGTGGT TTTTTAATTT ATTTAAAATT GGGGGGAGGG   960
```

325

```
CATGGGAAGG ATTTAACACC GATATATTGT TACCGCTGAA AATGAACTTT ATGAACCTTT   1020
TCCAAGTTGA TCTATCCAGT GACGTGGCCT GGTGGGCGTT TCTTCTTGTA CTTATGTGGT   1080
TTTTTGGCTT TTAATACAGA CATTTTCCTC CAAAAAAAAA AAAAAAAGG
```

5

SEQ ID NO:64 PDG8 Protein sequence
Protein Accession #:   CAB45781

10

```
1          11         21         31         41         51
|          |          |          |          |          |
GRRTGRLRPA AAPSAAAATA GAPTALPAYP AAEPPGPLWL QGEPLHFCCL DFSLEELQGE   60
PGWRLNRKPI ESTLVACFMT LVIVVWSVAA LIWPVPIIAG FLPNGMEQRR TTASTTAATP   120
AAVPAGTTAA AAAAAAAAAA AAVTSGVATK
```

SEQ ID NO:65 PDM1 DNA SEQUENCE

15

Nucleic Acid Accession #:   NM_006765
Coding sequence: 149-1195 (underlined sequences correspond to start and stop codons)

20

```
1          11         21         31         41         51
|          |          |          |          |          |
CGGCCGCGGC CCGGGTCCCT CGCAAAGCCG CTGCCATCCC GGAGGGCCCA GCCAGCGGGC   60
TCCCGGAGGC TGGCCGGGCA GGCGTGGTGC GCGGTAGGAG CTGGGCGCGC ACGGCTACCG   120
CGCGTGGAGG AGACACTGCC CTGCCGCGAT GGGGGCCCGG GGCGCTCCTT CACGCCGTAG   180
GCAAGCGGGG CGGCGGCTGC GGTACCTGCC CACCGGGAGC TTTCCCTTCC TTCTCCTGCT   240
GCTGCTGCTC TGCATCCAGC TCGGGGGAGG ACAGAAGAAA AAGGAGAATC TTTTAGCTGA   300
AAAAGTAGAG CAGCTGATGG AATGGAGTTC CAGACGCTCA ATCTTCCGAA TGAATGGTGA   360
TAAATTCCGA AAATTTATAA AGGCACCACC TCGAAACTAT TCCATGATTG TTATGTTCAC   420
TGCTCTTCAG CCTCAGCGGC AGTGTTCTGT GTGCAGGCAA GCTAATGAAG AATATCAAAT   480
ACTGGCGAAC TCCTGGCGCT ATTCATCTGC TTTTTGTAAC AAGCTCTTCT TCAGTATGGT   540
GGACTATGAT GAGGGGACAG ACGTTTTTCA GCAGCTCAAC ATGAACTCTG CTCCTACATT   600
CAYGCATTTW CCTCCAAAAG GCAGACCTAA GAGAGCTGAT ACTTTTGACC TCCAAAGAAT   660
TGGATTTGCA GCTGAGCAAC TAGCAAAGTG GATTGCTGAC AGAACGGATG TTCATATTCG   720
GGTTTTCAGA CCACCCAACT ACTCTGGTAC CATTGCTTTG GCCCTGTTAG TGTCGCTTGT   780
TGGAGGTTTG CTTTATTNGA GAAGGAACAA CTTGGAGTTC ATCTATAACA AGACTGGTTG   840
GGCCATGGTG TCTCTGTGTA TAGTCTTTGC TATGACTTCT GGCCAGATGT GGAACCATAT   900
CCGTGGACCT CCATATGCTC ATAAGAACCC ACACAATGGA CAAGTGAGCT ACATTCATGG   960
GAGCAGCCAG GCTCAGTTTG TGGCAGAATC ACACATTATT CTGGTACTGA ATGCCGCTAT   1020
CACCATGGGG ATGGTTCTTC TAAATGAAGC AGCAACTTCG AAAGGCGATG TTGGAAAAAG   1080
ACGGATAATT TGCCTAGTGG GATTGGGCCT GGTGGTCTTC TTCTTCAGTT TTCTACTTTC   1140
AATATTTCGT TCCAAGTACC ACGGCTATCC TTATAGTGAT CTGGACTTTG AGTGAGAAGA   1200
TGTGATTTGG ACCATGGCAC TTAAAAACTC TATAACCTCA GCTTTTTAAT TAAATGAAGC   1260
CAAGTGGGAT TTGCATAAAG TGAATGTTTA CCATGAAGAT AAACTGTTCC TGACTTTATA   1320
CTATTTTGAA TTCATTCATT TCATTGTGAT CAGCTAGCTT ATTCTTGTGT ACTTTTTTTA   1380
AACTGTGGGT TTTCCTAGTA AATTTAATTT ACAGAAATCA ATGGTAGCAT TTAGTAATCT   1440
ACAAAGGAAA TATCAAAGTG TTTTTCAAGC CTGTTATATY CAGTGTGTKC CACAGGATTG   1500
CAATAAATGA CAATGTAATT A
```

50

SEQ ID NO:66 PDM1 Protein sequence
Protein Accession #:   NP_006756

```
1          11         21         31         41         51
|          |          |          |          |          |
MGARGAPSRR RQAGRRLRYL PTGSFPFLLL LLLLCIQLGG GQKKKENLLA EKVEQLMEWS   60
SRRSIFRMNG DKFRKFIKAP PRNYSMIVMF TALQPQRQCS VCRQANEEYQ ILANSWRYSS   120
AFCNKLFFSM VDYDEGTDVF QQLNMNSAPT FXHXPPKGRP KRADTFDLQR IGFAAEQLAK   180
WIADRTDVHI RVFRPPNYSG TIALALLVSL VGGLLYXRRN NLEFIYNKTG WAMVSLCIVF   240
AMTSGQMWNH IRGPPYAHKN PHNGQVSYIH GSSQAQFVAE SHIILVLNAA ITMGMVLLNE   300
AATSKGDVGK RRIICLVGLG LVVFFFSFLL SIFRSKYHGY PYSDLDFE
```

60

SEQ ID NO:67 PDM2 DNA SEQUENCE

Nucleic Acid Accession #:   NM_000947
Coding sequence: 88-1617 (underlined sequences correspond to start and stop codons)

65

```
1          11         21         31         41         51
|          |          |          |          |          |
GGTTTCATAT GAACTCTCCC GCCACCCGGG AACAGCTGGC TGCCACCGTT TGTGTTTTCC   60
GAGTTTGTAT TCTTGCAGGT GACCAAGATG GAGTTTTCTG GAAGAAAGCG GAGGAAGCTG   120
AGGTTGGCAG GTGACCAGAG GAATGCTTCC TACCCTCATT GCCTTCAGTT TTACTTGCAG   180
CCACCTTCTG AAAACATATC TTTAACAGAA TTTGAAAACT TGGCTATTGA TAGAGTTAAA   240
TTGTTAAAAT CAGTTGAAAA TCTTGGAGTG AGCTATGTGA AAGGAACTGA ACAATACCAG   300
AGTAAGTTGG AGAGTGAGCT TCGGAAGCTC AAGTTTTCCT ACAGAGAGAA GCTAGAAGAT   360
GAATATGAAC CACGAAGAAG AGATCATATT TCTCATTTTA TTTTGCGGCT TGCTTATTGA   420
CAGTCTGAAG AACTTAGACG CTGGTTCATT CAACAAGAAA TGGATCTCCT TCGATTTAGA   480
TTTAGTATTT TACCCAAGGA TAAAATTCAG GATTTCTTAA AGGATAGCCA ATTGCAGTTT   540
GAGGCTATAA GTGATGAAGA GAAGACTCTT CGAGAACAGG AGATTGTTGC CTCATCACCA   600
AGTTTAAGTG GACTTAAGTT GGGGTTCGAG TCCATTTATA AGATCCCTTT TGCTGATGCT   660
CTGGATTTGT TTCGAGGAAG GAAAGTCTAT TTGGAAGATG GCTTTGCTTA CGTACCACTT   720
```

```
AAGGACATTG TGGCAATCAT CCTGAATGAA TTTAGAGCCA AACTGTCCAA GGCTTTGGCA    780
TTAACAGCCA GGTCCTTGCC TGCTGTGCAG TCTGATGAAA GACTTCAGCC TCTGCTCAAT    840
CACCTCAGTC ATTCCTACAC TGGCCAAGAT TACAGTACCC AGGGAAATGT TGGGAAGATT    900
TCTTTAGATC AGATTGATTT GCTTTCTACC AAATCCTTCC CACCTTGCAT GCCGTCAGTTA   960
CATAAAGCCT TGCGGGAAAA TCACCATCTT CGTCATGGAG GCCGAATGCA GTATGGCCTA   1020
TTTCTGAAGG GCATTGGTTT AACTTTGGAA CAGGCATTGC AGTTCTGGAA GCAAGAATTT   1080
ATCAAAGGAA AGATGGATCC AGACAAGTTT GATAAAGGTT ACTCTTACAA CATCCGTCAC   1140
AGCTTTGGAA AGGAAGGCAA GAGGACAGAC TATACACCTT TCAGTTGCCT GAAGATTATT   1200
CTGTCCAATC CACCAAGCCA AGGGGATTAT CATGGGTGCC CATTCCGTCA CAGTGATCCA   1260
GAGCTGCTGA AGCAAAAGTT GCAGTCATAC AAGATCTCTC CTGGAGGGAT AAGCCAGATT   1320
TTGGATTTAG TAAAGGGGAC ACATTACCAG GTAGCCTGTC AAAAATACTT TGAGATGATA   1380
CACAATGTGG ATGATTGTGG CTTTTCTTTG AATCATCCTA ATCAGTTCTT TTGTGAGAGC   1440
CAACGTATTC TAAATGGTGG TAAAGACATA AAGAAGGAAC CTATCCAACC AGAAACTCCT   1500
CAACCCAAAC CAAGTGTCCA GAAAACCAAG GATGCATCAT CTGCTCTGGC CTCTTTAAAT   1560
TCCTCTCTGG AAATGGATAT GGAAGGACTA GAAGATTACT TTAGTGAAGA TTCTTAGGCA   1620
GTTTTATAAC CCTTTTTCCT CAATAGCCTG TTTCCTGTTT TTAAGATTTT GCCTTTGTTG   1680
TTGAAAAAGG GTTTCACTGT CACCAAGGCT TAGTGCAGTG ACACAATTAC AGCTGATTGC   1740
AGCCTTGACC TTCCCAGCTC AAGTGATCCT CCTACCTCAG CCTCCCAAGT AGTTAGGACA   1800
CACAGGTGTG CACCTCATAT CCAGATAATT TTTTTCAATT TTTTTTTGTA GAGGTGGGGG   1860
GTCTCCCTAT GTTGCCCAGG CAGATCTCAG ACTCCTGGGC TCAAGCGATC CTCACACCTC   1920
AGCGTCCCAG AGTGCTGGGA TTACAGTTGT GAGCCACTGT GCCTGGCCTT TTTTTTTTTT   1980
TAACCTTTTC GTTTAACTTC TCTCTTCACT GCATCCCAAT CCATCTACAG GCATGCACAC   2040
TTATTAGGAA AGGAGGTTTG AGGTAACAAC AGAGACTTTC ACTATATTTT GCTTTGACAG   2100
AAGGAAAGAG GAGGAGTTTC TATTAAAATC TGTCACTTGA GTGATGTCAT TTAAGTCCTA   2160
TTTTTAGGAGA TAAAAACAGC TTTGGGGACT GGTTAAAGTC CCCCAGAAAC TACAATAAAG   2220
AACAACTTTT GTTTTAACTC TTAATCACTT TGTAATTTTG ACTCAATCCT TTTCTGGACC   2280
ATTTTTGTTA ATAAATATCA AAGTGT
```

SEQ ID NO:68 PDM2 Protein sequence:
Protein Accession #:   NP_000938

```
1          11         21         31         41         51
|          |          |          |          |          |
MEFSGRKRRK LRLAGDQRNA SYPHCLQFYL QPPSENISLT EFENLAIDRV KLLKSVENLG    60
VSYVKGTEQY QSKLESELRK LKFSYREKLE DEYEPRRRDH ISHFILRLAY CQSEELRRWF   120
IQQEMDLLRF RFSILPKDKI QDFLKDSQLQ FEAISDEEKT LREQEIVASS PSLSGLKLGF   180
ESIYKIPFAD ALDLFRGRKV YLEDGFAYVP LKDIVAIILN EFRAKLSKAL ALTARSLPAV   240
QSDERLQPLL NHLSHSYTGQ DYSTQGNVGK ISLDQIDLLS TKSFPPCMRQ LHKALRENHH   300
LRHGGRMQYG LFLKGIGLTL EQALQFWKQE FIKGKMDPDK FDKGYSYNIR HSFGKEGKRT   360
DYTPFSCLKI ILSNPPSQGD YHGCPFRHSD PELLKQKLQS YKISPGGISQ ILDLVKGTHY   420
QVACQKYFEM IHNVDDCGFS LNHPNQFFCE SQRILNGGKD IKKEPIQPET PQPKPSVQKT   480
KDASSALASL NSSLEMDMEG LEDYFSEDS
```

SEQ ID NO:69 PDM3 DNA SEQUENCE

Nucleic Acid Accession #:   NM_024840
Coding sequence: 108-491 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
AATTCATACA GGAGAGAAGT CATATATATG CAGTGATTGT GGAAAAGGCT TCATCAAGAA    60
GTCTCGGCTC ATTAATCATC AGAGAGTTCA TACAGGAGAG AAACCACATG GATGCAGCCT   120
GTGTGGGAAG GCCTTCTCCA AAAGGTCCAG GCTCACTGAA CACCAGAGAA CTCATACAGG   180
AGAGAAGCCC TATGAATGCA CTGAATGTGA CAAAGCATTC CGCTGGAAAT CACAGCTCAA   240
TGCACATCAG AAAGCTCACA CAGGAGAGAA GTCATATATA TGCCGTGATT GTGGAAAAGG   300
CTTCATTCAG AAGGGAAATC TCATTGTACA TCAGCGAATT CATACTGGAG AAAAACCCTA   360
TATATGCAAT GAATGTGGAA AAGGCTTCAT CCAAAAGGGC AACCTCCTTA TTCATCGACG   420
TACTCACACT GGAGAGAAAC CCTATGAATG CAATGAATGT GGGAAAGGCT TCAGCCAGAA   480
GACATGTTTA ATATCCCATC AGAGATTTCA CACAGGAAAG ACACCCTTTG TATGTACTGA   540
GTGTGGAAAA TCCTGCTCAC ACAAGTCAGG TCTCATTAAC CACCAGAGAA TTCACACAGG   600
AGAGAAACCC TATACATGCA GTGACTGTGG GAAAGCTTTC AGAGATAAAT CATGTCTCAA   660
CAGACATCGG AGAACTCATA CAGGGGAGAG ACCGTATGGA TGCTCTGATT GTGGGAAAGC   720
TTTCTCCCAC TTGTCATGCC TTGTTTATCA TAAGGGAATG CTGCATGCAA GAGAGAAATG   780
TGTAGGTTCA GTCAAATTGG AAAATCCTTG CTCAGAGAGT CATAGCTTAT CACATACACG   840
TGATCTCATA CAGGATAAAG ACTCTGTTAA CATGGTGACT CTGCAGATGC CTTCTGTGGC   900
AGCTCAGACC TCATTAACTA ACAGTGCGTT CCAAGCAGAG AGCAAAGTAG CCATTGTGAG   960
CCAGCCTGTT GCCAGAAGTT CAGTCTCAGC AGATAGTAGA ATTTGCACAG AATAAAAACC   1020
ATATGAATGC AGTGAATGTG GTAGTGCTTT CAGTGATCAA TTACATCATA TGTCACAAAA   1080
AACACAGAGG AACAAACTGA TATATTCAAG GTGGAAAGCC CTTGAATAAA ACCTTATGGC   1140
TAATAAGCAT ATACTCAGAG AAAAATAGTA TGAAGTGGAG ACTGGGAAAT TCTTTTATGG   1200
GAAGATAGAT CTTCTCATCA GTGACCATAG ATCACATCTT CAGTGAGCTT ATAGTTGGTA   1260
GAAATATAAT GATCATGGAA AAGTCCTTGT TCAGAAACAG TACGCCAGTA GGTATCAGGG   1320
GGTTTACACA GGAGAGAAAC TTTTGGAAGA CCTTTGAAGG CTATGAATGT GGCAGGGTTG   1380
CTAGTGGTAC ATTCTGCCTT ATCCTCAGAG GGAATCATAT AGAAATAAAA CTATGAAAAT   1440
GTAACTAGAA CATCTTCATC AAAATATGAA AGAACACACG AAGCAAATAA GCCCTGTGAA   1500
AAGGAGTATT TTAGAGATTT CGATCAGAAA TCTAACATCA TTATATGGCA GATAATATAC   1560
AGGATGTGTA TTTTAGGACA ATATACCTTG AATCACTAGT TGATATGTCA ATGACTAATT   1620
AAAAGGGGTT GTCAGTGTTA CACATCATTG GTTAAATTTA TAGCACAATG TACCTCTTCC   1680
CCCTTTTTTG ATAAGAGTCT TCTATTCCCA ACCAAGATCA TTATATGATT AGCTCTTGTG   1740
TTTCTTTGAT TCCAAATTTC TTCACTTGTT ATTTCAGACT ACTGAGAAGA TTCAAAAGGA   1800
```

327

```
AAAATGTATT TAATTTAATA ATGTAACACA ACAAGTTTGG ATGTGTTTAA CTTTATAAAT  1860
AATCACCCCA GAGGAATGAA GTTCAAAACT TGTGAATAAC C
```

5 SEQ ID NO:70 PDM3 Protein sequence:
Protein Accession #:   NP_079116

```
1          11         21         31         41         51
|          |          |          |          |          |
MDAACVGRPS PKGPGSLNTR ELIQERSPMN ALNVTKHSAG NHSSMHIRKL TQERSHIYAV   60
IVEKASFRRE ISLYISEFIL EKNPIYAMNV EKASSKRATS LFIDVLTLER NPMNAMNVGK   120
ASARRHV
```

SEQ ID NO:71 PDM8 DNA SEQUENCE

Nucleic Acid Accession #: NM_018455
Coding sequence: 341-955 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
AATTTCGGCA CGGGGGGGAG GCACAGTGAG TCCACTGGGG CACGGCAGCG TCTAAGCCAC   60
AAGCCGACTG ACATAAGCCA GGTCCTAACG GAGCCTATGT GTAAGTCCAC TACTGGTGCA   120
AGGTTGCACA CTTCTAAGAA GAGCGGCGTG GGGGGCTCGG CGACCTTCGC TTCAGTCGCT   180
CCCCCGTGCA GTCCCCTGTG CCCAAGACAC AGCCTGATGC TTGTGCTCCG GTGGGCGGAC   240
TTGGAGGCGG CGGGAACTGC AATTGGTGGC TTTGAAGGGC GGCGAGCGGG AACAGCTCTT   300
GAGGAGTGAG ACTGCAGGAG ATGTGGGCCG TGCCAAAGAG ATGGATGAGA CTGTTGCTGA   360
GTTCATCAAG AGGACCATCT TGAAAATCCC CATGAATGAA CTGACAACAA TCCTGAAGGC   420
CTGGGATTTT TTGTCTGAAA ATCAACTGCA GACTGTAAAT TTCCGACAGA GAAAGGAATC   480
TGTAGTTCAG CACTTGATCC ATCTGTGTGA GGAAAAGCGT GCAAGTATCA GTGATGCTGC   540
CCTGTTAGAC ATCATTTATA TGCAATTTCA TCAGCACCAG AAAGTTTGGG ATGTTTTTCA   600
GATGAGTAAA GGACCAGGTG AAGATGTTGA CCTTTTTGAT ATGAAACAAT TTAAAAATTC   660
GTTCAAGAAA ATTCTTCAGA GAGCATTAAA AAATGTGACA GTCAGCTTCA GAGAAACTGA   720
GGAGAATGCA GTCTGGATTG CAATTGCCTG GGGAACACAG TACACAAAGC CAAACCAGTA   780
CAAACCTACC TACGTGGTGT ACTACTCCCA GACTCCGTAC GCCTTCACGT CCTCCTCCAT   840
GCTGAGGCGC AATACACCGC TTCTGGGTCA GGAGTTAGAA GCTACTGGGA AAATCTACCT   900
CCGACAAGAG GAGATCATTT TAGATATTAC CGAAATGAAG AAAGCTTGCA ATTAGTGAAC   960
ATGAAAGGAA AATAAAAATT CCTCACAGTC AAAAAAAAAA AAAAA
```

SEQ ID NO:72 PDM8 Protein sequence:
Protein Accession #:   NP_060925

```
1          11         21         31         41         51
|          |          |          |          |          |
MDETVAEFIK RTILKIPMNE LTTILKAWDF LSENQLQTVN FRQRKESVVQ HLIHLCEEKR   60
ASISDAALLD IIYMQFHQHQ KVWDVFQMSK GPGEDVDLFD MKQFKNSFKK ILQRALKNVT   120
VSFRETEENA VWIRIAWGTQ YTKPNQYKPT YVVYYSQTPY AFTSSSMLRR NTPLLGQELE   180
ATGKIYLRQE EIILDITEMK KACN
```

SEQ ID NO:73 PDM9 DNA SEQUENCE

Nucleic Acid Accession #: NM_016192
Coding sequence: 1-1125 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGGTGCTGT GGGAGTCCCC GCGGCAGTGC AGCAGCTGGA CACTTTGCGA GGGCTTTTGC   60
TGGCTGCTGC TGCTGCCCGT CATGCTACTC ATCGTAGCCC GCCCGGTGAA GCTCGCTGCT   120
TTCCCTACCT CCTTAAGTGA CTGCCAAACG CCCACCGGCT GGAATTGCTC TGGTTATGAT   180
GACAGAGAAA ATGATCTCTT CCTCTGTGAC ACCAACACCT GTAAATTTGA TGGGGAATGT   240
TTAAGAATTG GAGCACTGT GACTTGCGTC TGTCAGTTCA AGTGCAACAA TGACTATGTG   300
CCTGTCTGTG GCTCCAATGG GGAGAGCTAC CAGAATGAGT GTTACCTGCG ACAGGCTGCA   360
TGCAAACAGC AGAGTGAGAT ACTTGTGGTG TCAGAAGGAT CATGTGCCAC AGATGCAGGA   420
TCAGGATCTG GAGATGGAGT CCATGAAGGC TCTGGAGAAA CTAGTCAAAA GGAGACATCC   480
ACCTGTGATA TTTGCCAGTT TGGTGCAGAA TGTGACGAAG ATGCCGAGGA TGTCTGGTGT   540
GTGTGTAATA TTGACTGTTC TCAAACCAAC TTCAATCCCC TCTGCGCTTC TGATGGGAAA   600
TCTTATGATA ATGCATGCCA AATCAAAGAA GCATCGTGTC AGAAACAGGA GAAAATTGAA   660
GTCATGTCTT TGGGTCGATG TCAAGATAAC ACAACTACAA CTACTAAGTC TGAAGATGGG   720
CATTATGCAA GAACAGATTA TGCAGAGAAT GCTAACAAAT TAGAAGAAAG TGCCAGAGAA   780
CACCACATAC CTTGTCCGGA ACATTACAAT GGCTTCTGCA TGCATGGGAA GTGTGAGCAT   840
TCTATCAATA TGCAGGAGCC ATCTTGCAGG TGTGATGCTG GTTATACTGG ACAACACTGT   900
GAAAAAAAGG ACTACAGTGT TCTATACGTT GTTCCCGGTC CTGTACGATT TCAGTATGTC   960
TTAATCGCAG CTGTGATTGG AACAATTCAG ATTGCTGTCA TCTGTGTGGT GGTCCTCTGC  1020
ATCACAAGGA AATGCCCCAG AAGCAACAGA ATTCACGAC AGAAGCAAAA TACAGGGCAC  1080
TACAGTTCAG ACAATACAAC AAGAGCGTCC ACGAGGTTAA TCTGA
```

328

SEQ ID NO:74 PDM9 Protein sequence:
Protein Accession #:   NP_057276

```
              1         11         21         31         41         51
              |          |          |          |          |          |
     1 MVLWESPRQC SSWTLCEGFC WLLLLPVMLL IVARPVKLAA FPTSLSDCQT PTGWNCSGYD    60
    61 DRENDLFLCD TNTCKFDGEC LRIGDTVTCV CQFKCNNDYV PVCGSNGESY QNECYLRQAA   120
   121 CKQQSEILVV SEGSCATDAG SGSGDGVHEG SGETSQKETS TCDICQFGAE CDEDAEDVWC   180
   181 VCNIDCSQTN FNPLCASDGK SYDNACQIKE ASCQKQEKIE VMSLGRCQDN TTTTTKSEDG   240
   241 HYARTDYAEN ANKLEESARE HHIPCPEHYN GFCMHGKCEH SINMQEPSCR CDAGYTGQHC   300
   301 EKKDYSVLYV VPGPVRFQYV LIAAVIGTIQ IAVICVVVLC ITRKCPRSNR IHRQKQNTGH   360
   361 YSSDNTTRAS TRLI
```

SEQ ID NO:75 PDO1 DNA SEQUENCE

Nucleic Acid Accession #:   NM_014324
Coding sequence: 89-1237 (underlined sequences correspond to start and stop codons)

```
         1         11         21         31         41         51
         |          |          |          |          |          |
GGCGCCGGGA TTGGGAGGGC TTCTTGCAGG CTGCTGGGCT GGGGCTAAGG GCTGCTCAGT    60
TTCCTTCAGC GGGGCACTGG GAAGCGCCAT GGCACTGCAG GGCATCTCGG TCGTGGAGCT   120
GTCCGGCCTG GCCCCGGGCC GTNTCTGTGC TATGGTCCTG GCTGACTTCG GGGCGCGTGT   180
GGTACGCGTG GACCGGCCCG GCTCCCGCTA CGACGTGAGC CGCTTGGGCC GGGGCAAGCG   240
CTCGCTAGTG CTGGACCTGA AGCAGCCGCG GGAGCCGCGT GCTGCGGCGT CTGTGCAAGC   300
GGTCGGATGT GCTGCTGGAG CCCTTCCGCC GCGGTGTCAT GGAGAAACTC CAGCTGGGCC   360
CAGAGATTCT GCAGCGGGAA AATCCAAGGC TTATTTATGC CAGGCTGAGT GGATTTGGCC   420
AGTTCAGGAA AGCTTCTGCC GGTTAGCTGG CCACGATATC AACTATTTGG CTTTGTCAGG   480
TGTTCTCTCA AAAATTGGCA GAAGTGGTGA GAATCCGTAT GCCCCGCTGA ATCTCGTGGC   540
TGACTTTGCT GGTGGTGGCC TTATGTGTGC ACTGGGCATT ATAATGGCTC TTTTTGACCG   600
CACACGCACT GACAAGGGTC AGGTCATTGA TGCAAATATG GTGGAAGGAA CAGCATATTT   660
AAGTTCTTTT CTGTGGAAAA CTCAGAAATC GAGTCTGTGG GAAGCACCTC GAGGACAGAA   720
CATGTTGGAT GGTGGAGCAC CTTTCTATAC GACTTACAGG ACAGCAGATG GGGAATTCAT   780
GGCTGTTGGA GCAATAGAAC CCCAGTTCTA CGAGCTGCTG ATCAAAGGAC TTGGACTAAA   840
GTCTGATGAA CTTCCCAATC AGATGAGCAC GGATGATTGG CCAGAAATGA AGAAGAAGTT   900
TGCAGATGTA TTTGCAAAGA AGACGAAGGC AGAGTGGTGT CAAATCTTTG ACGGCACAGA   960
TGCCTGTGTG ACTCCGGTTC TGACTTTTGA GGAGGTTGTT CATCATGATC ACAACAAGGA  1020
ACGGGGCTCG TTTATCACCA GTGAGGAGCA GGACGTGAGC CCCCGCCTTG CACCTCTGCT  1080
GTTAAACACC CCAGCCATCC CTTCTTCCAA AGGGGATCCT TTCATAGGAG AACACACTGA  1140
GGAGATACTT GAAGAATTTG GATTCAGCCG AGAAGAGATT TATCAGCTTA ACTCAGATAA  1200
AATCATTGAA AGTAATAAGG TAAAAGCTAG TCTCTAACTT CCAGGCCCAC GGCTCAAGTG  1260
AATTTGAATA CTGCATTTAC AGTGTAGAGT AACACATAAC ATTGTATGCA TGGAAACATG  1320
GAGGAACAGT ATTACAGTGT CCTACCACTG TAATCAAGAA AAGAATTACA GACTCTGATT  1380
CTACAGTGAT GATTGAATTC TAAAAATGGT TATCATTAGG GCTTTTGATT TATAAAACTT  1440
TGGGTACTTA TACTAAATTA TGGTAGTTAT TCTGCCTTCC AGTTTGCTTG ATATATTTGT  1500
TGATATTAAG ATTCTTGACT TATATTTTGA ATGGGTTCTA GTGAAAAAGG AATGATATAT  1560
TCTTGAAGAC ATCGATATAC ATTTATTTAC ACTCTTGATT CTACAATGTA GAAAATGAGG  1620
AAATGCCACA AATTGTATGG TGATAAAAGT CACGTGAAAC AGAGTGATTG GTTGCATCCA  1680
GGCCTTTTGT CTTGGTGTTC ATGATCTCCC TCTAAGCACA TTCCAAACTT TAGCAACAGT  1740
TATCACACTT TGTAATTTGC AAAGAAAAGT TTCACCTGTA TTGAATCAGA ATGCCTTCAA  1800
CTGAAAAAAA CATATCCAAA ATAATGAGGA AATGTGTTGG CTCACTACGT AGAGTCCAGA  1860
GGGACAGTCA GTTTTAGGGT TGCCTGTATC CAGTAACTCG GGGCCTGTTT CCCCGTGGGT  1920
CTCTGGGCTG TCAGCTTTCC TTTCTCCATG TGTTTGATTT CTCCTCAGGC TGGTAGCAAG  1980
TTCTGGATCT TATACCCAAC ACACAGCAAC ATCCAGAAAT AAAGATCTCA GGACCCCCCA  2040
AAAAAAAAA AAAAAAAAA AAAAAAA
```

SEQ ID NO:76 PDO1 Protein sequence:
Protein Accession #:   NP_055139

```
              1         11         21         31         41         51
              |          |          |          |          |          |
     1 MALQGISVVE LSGLAPGRXC AMVLADFGAR VVRVDRPGSR YDVSRLGRGK RSLVLDLKQP    60
    61 REPRAAASVQ AVGCAAGALP PRCHGETPAG PRDSAAGKSK AYLCQAEWIW PVQESFCRLA   120
   121 GHDINYLALS GVLSKIGRSG ENPYAPLNLV ADFAGGGLMC ALGIIMALFD RTRTDKGQVI   180
   181 DANMVEGTAY LSSFLWKTQK SSLWEAPRGQ NMLDGGAPFY TTYRTADGEF MAVGAIEPQF   240
   241 YELLIKGLGL KSDELPNQMS TDDWPEMKKK FADVFAKKTK AEWCQIFDGT DACVTPVLTF   300
   301 EEVVHHDHNK ERGSFITSEE QDVSPRLAPL LLNTPAIPSS KGDPFIGEHT EEILEEFGFS   360
   361 REEIYQLNSD KIIESNKVKA SL
```

SEQ ID NO:77 PDO3 DNA SEQUENCE

Nucleic Acid Accession #:   AB028951
Coding sequence: 97-1128 (underlined sequences correspond to start and stop codons)

```
         1         11         21         31         41         51
         |          |          |          |          |          |
GTTAAATCCT TACTTTACCA GATTCTTGAT GGTATCCATT ACCTCCATGC AAATTGGGTG    60
CTTCACAGAG ACTTGAAACC AGCAAATATC CTAGTAATGG GAGAAGGTCC TGAGAGGGGG   120
AGAGTCAAAA TAGCTGACAT GGGTTTTGCC AGATTATTCA ATTCTCCTCT AAAGCCACTA   180
GCAGATTTGG ATCCAGTAGT TGTGACATTT TGGTATCGGG CTCCAGAACT TTTGCTTGGT   240
GCAAGGCATT ATACAAAGGC CATTGATATA TGGGCAATAG GTTGTATATT TGCTGAATTG   300
TTGACTTCGG AACCTATTTT TCACTGTCGT CAGGAAGATA TAAAAACAAG CAATCCCTTT   360
```

```
        CATCATGATC AACTGGATCG GATATTTAGT GTCATGGGGT TTCCTGCAGA TAAAGACTGG   420
        GAAGATATTA GAAAGATGCC AGAATATCCC ACACTTCAAA AAGACTTTAG AAGAACAACG   480
        TATGCCAACA GTAGCCTCAT AAAGTACATG GAGAAACACA AGGTCAAGCC TGACAGCAAA   540
        GTGTTCCTCT TGCTTCAGAA ACTCCTGACC ATGGATCCAA CCAAGAGAAT TACCTCGGAG   600
   5    CAAGCTCTGC AGGATCCCTA TTTTCAGGAG GACCCTTTGC CAACATTAGA TGTATTTGCC   660
        GGCTGCCAGA TTCCATACCC CAAAACGAGAA TTCCTTAATG AAGATGATCC TGAAGAAAAA   720
        GGTGACAAGA ATCAGCAACA GCAGCAAGAC CAGCATCAGC AGCCCACAGC CCCTCCACAG   780
        CAGGCAGCAG CCCCTCCACA GGCGCCCCCA CCACAGCAGA ACAGCACCCA GACCAACGGG   840
        ACCGCAGGTG GGGCTGGGGC CGGGGTCGGG GGCACCGGAG CAGGGTTGCA GCACAGCCAG   900
  10    GACTCCAGCC TGAACCAGGT GCCTCCAAAC AAGAAGCCAC GGCTAGGGCC TTCAGGCGCA   960
        AACTCAGGTG GACCTGTGAT GCCCTCGGAT TATCAGCACT CCAGTTCTCG CCTGAATTAC  1020
        CAAAGCAGCG TTCAGGGATC CTCTCAGTCC CAGAGCACAC TTGGCTACTC TTCCTCGTCT  1080
        CAGCAGAGCT CACAGTACCA CCCATCTCAC CAGGCCCACC GGTACTGACC AGCTCCCGTT  1140
        GGGCCAGGCC AGCCCAGCCC AGAGCACAGG CTCCAGCAAT ATGTCTGCAT TGAAAAGAAC  1200
  15    CAAAAAAATG CAAACTATGA TGCCATTTAA AACTCATACA CATGGGAGGA AAACCTTATA  1260
        TACTGAGCAT TGTGCAGGAC TGATAGCTCT TCTTTATTGA CTTAAAGAAG ATTCTTGTGA  1320
        AGTTTCCCCA GCACCCCTTC CCTGCATGTG TTCCATTGTG ACTTCTCTGA TAAAGCGTCT  1380
        GATCTAATCC CAGCACTTCT GTAACCTTCA GCATTTCTTT GAAGGATTTC CTGGTGCACC  1440
        TTTCTCATGC TGTAGCAATC ACTATGGTTT ATCTTTTCAA AGCTCTTTTA ATAGGATTTT  1500
  20    AATGTTTTAG AAACAGGATT CCAGTGGTGT ATAGTTTTAT ACTTCATGAA CTGATTTAGC  1560
        AACACAGGTA AAAATGCACC TTTTAAAGCA CTACGTTTTC ACAGACAATA ACTGTTCTGC  1620
        TCATGGAAGT CTTAAACAGA AACTGTTACT GTCCCAAAGT ACTTTACTAT TACGTTCGTA  1680
        TTTATCTAGT TTCAGGGAAG GTCTAATAAA AAGACAAGCG GTGGGACAGA GGGAACCTAC  1740
        AACCAAAAAC TGCCTAGATC TTTGCAGTTA TGTGCTTTAT GCCACGAAGA ACTGAAGTAT  1800
  25    GTGGTAATTT TTATAGAATC ATTCATATGG AACTGAGTTC CCAGCATCAT CTTATTCTGA  1860
        ATAGCATTCA GTAATTAAGA ATTACAATTT TAACCTTCAT GTAGCATAAGT CTACCTTAAA  1920
        AAGGGTTTCA AGAGCTTTGT ACAGTCTCGA TGGCCCACAC CAAAACGCTG AAGAGAGTAA  1980
        CAACTGCACT AGGATTCTG TAAGGAGTAA TTTTGATCAA AAGACGTGTT ACTTCCCTTT  2040
        GAAGGAAAAG TTTTTAGTGT GTATTGTACA TAAAGTCGGC TTCTCTAAAG AACCATTGGT  2100
  30    TTCTTCACAT CTGGGTCTGC GTGAGTAACT TTCTTGCATA ATCAAGGTTA CTCAAGTAGA  2160
        AGCCTGAAAA TTAATCTGCT TTTAAAATAA AGAGCAGTGT TCTCCATTCG TATTTGTATT  2220
        AGATATAGAG TGACTATTTT TAAAGCATGT TAAAAATTTA GGTTTTATTC ATGTTTAAAG  2280
        TATGTATTAT GTATGCATAA TTTTGCTGTT GTTACTGAAA CTTAATTCTA TCAAGAATCT  2340
        TTTTCATTGC ACTGAATGAT TTCTTTTGCC CCTAGGAGAA AACTTAATAA TTGTGCCTAA  2400
  35    AAACTATGGG CGGATAGTAT AAGACTATAC TAGACAAAGT GAATATTTGC ATTTCCATTA  2460
        TCTATGAATT AGTGGCTGAG TTCTTTCTTA GCTGCTTTAA GGAGCCCCTC ACTCCCCAGA  2520
        GTCAAAAGGA AATGTAAAAA CTTAGAGCTC CCATTGTAAT GTAAGGGGCA AGAAATTTGT  2580
        GTTCTTCTGA ATGCTACTAG CAGCACCAGC CTTGTTTTAA ATGTTTTCTT GAGCTAGAAG  2640
        AAATAGCTGA TTATTGTATA TGCAAATTAC ATGCATTTTT AAAAACTATT CTTTCTGAAC  2700
  40    TTATCTACCT GGTTATGATA CTGTGGGTCC ATACACAAGT AAAATAAGAT TAGACAGAAG  2760
        CCAGTATACA TTTTGCACTA TTGATGTGAT ACTGTAGCCA GCCAGGACCT TACTGATCTC  2820
        AGCATAATAA TGCTCACTAA TAATGAAGTC TGCATAGTGA CACTCATCAA GACTGAAGAT  2880
        GAAGCAGGTT ACGTGCTCCA TTGGAAGGAG TTTCTGATAG TCTCCTGCTG TTTTACCCCT  2940
        TCCATTTTTT AAAATAAGAA ATTAGCAGCC CTCTGCATAA TGTAGCTGCC TATATGCAGT  3000
  45    TTTATCCTGT GCCCTAAAGC CTCACTGTCC AGAGCTGTTG GTCATCAGAT GCTTATTGCA  3060
        CCCTCACCAT GTGCCTGGTG CCCTGCTGGG TAGAGAACAC AGAGGACAGG GCATACTTCT  3120
        TGTCCTTAAG GAGCTTGTGA TCTGTGACAG TAAGCCCTCC TGGGATGTCT GTGCCATGTG  3180
        ATTGACTTAC AAGTGAAACT GTCTTATAAT ATGAAGGTCT TTTTGTTTAC TTCTAAACCC  3240
        ACTTGGGTAG TTACTATCCC CAAATCTGTT CTGTAAATAA TATTATGGAA GGGTTTCTAT  3300
  50    GTCAGTCTAC CTTAGAGAAA GCCAGTGATT CAATATCACA AAAGGCATTG ACGTATCTTT  3360
        GAAATGTTCA CAGCAGCCTT TTAACAACAA CTGGGTGGTC CTTGTAGGCA GAACATACTC  3420
        TCCTAAGTGG TTGTAGGAAA TTGCAAGGAA AATAGAAGGT CTGTTCTTGC TCTCAAGGAG  3480
        GTTACCTTTA ATAAAAGAAG ACAAACCCAG ATAGAGATATGT AAACCAAAAT ACTATGCCCC  3540
        TTAATACTTT ATAAGCAGCA TTGTTAAAATA GTTCTTACGC TTATACATTC ACAGAACTAC  3600
  55    CCTGTTTTCC TTGTATATAA TGACTTTTGC TGGCAGAACT GAAATATAAA CTGTAAGGGG  3660
        ATTTCGTCAG TTGCTCCCAG TATACAATAT CCTCCAGGAC ATAGCCAGAA ATCTCCATTC  3720
        CACACATGAC TGAGTTCCTA TCCCTGCACT GGTACTGGCT CTTTTCTCCT CTTTCCTTGC  3780
        CTCAGGGTTC GTGCTACCCA CTGATTCCCT TTACCCTTAG TAATAATTTT GGATCATTTT  3840
        CTTTCCTTTA AAGGGGAACA AAGCCTTTTT TTTTTTTGAG ACGGAGTGTT GCTCTGTCAC  3900
  60    CCAAGCTGGA GTGCAGTGGC ACGATCTTGG CTCACTCCAA CCTCCACCTT CCAGGTTCAA  3960
        GTGATTCTCC TGCCTCAGCC TCCCGAGTAG CTGGGACTAC GGGCACGCAC CACCACGTCT  4020
        GGCTAATTTT TGTATTTTTA GTAGAGATGG GGTTTCACCC TATTGGTCAG GCTGGTCTTG  4080
        AATTCCTCAC CTCAGGTCAT CCGCCTGTCT CGGCCTCCCG AAGTGCTGGG ATTATAGGTG  4140
        TGAGCCACCG CACCCAGTTG GGAACAAAGC CTTTTTAACA CACGTAAGGG CCCTCAAACC  4200
  65    GTGGGACCTC TAAGGAGACC TTTGAAGCTT TTTGAGGGCA AACTTTACCT TTGTGGTCTG  4260
        CAAATGATGG CATTTCTCTT TGAAATTTAT TAGATACTGT TATGTCCCCC AAGGGTACAG  4320
        GAGGGGCATC CCTCAGCCTA TGGGAACACC CAAACTAGGA GGGGTTATTG ACAGGAAGGA  4380
        ATGAATCCAA GTGAAGGCTT TCTGCTCTTC GTGTTACAAA CCAGTTTCAG AGTTAGCTTT  4440
        CTGGGGAGGT GTGTGTTTGT GAAAGGAATT CAAGTGTTGC AGGACAGATG AGCTCAAGGT  4500
  70    AAGGTAGCTT TGGCAGCAGG GCTGATACTA TGAGGCTGAA ACAATCCTTG TGATGAAGTA  4560
        GATCATGCAG TGACATACAA AGACCAAGGA TTATGTATAT TTTTATATCT CTGTTGGTTTT  4620
        GAAACTTTAG TACTTAGAAT TTTTGGCCTTC TGCACTACTC TTTTGCTCTT ACGAACATAA  4680
        TGGACTCTTA AGAATGGAAA GGGATGACAT TTACCTATGT GTGCTGCCTC ATTCCTGGTG  4740
        AAGCAACTGC TACTTGTTCT CTATGCCTCT AAAATGATGC TGTTTTCTCT GCTAAAGGTA  4800
  75    AAAGAAAAGA AAAAAATAGT TGGAAAATAA GACATGCAAC TTGATGTGCT TTTGAGTAAA  4860
        TTTATGCAGC AGAAACTATA CAATGAAGGA AGAATTCTAT GGAAATTACA AATCCAAAAC  4920
        TCTATGATGA TGTCTTCCTA GGGAGTAGAG AAAGGCAGTG AAATGGCAGT TAGACCAACA  4980
        GAGGCTTGAA GGATTCAAGT ACAAGTAATA TTTTGTATAA AACATAGCAG TTTAGGTCCC  5040
        CATAATCCTC AAAAATAGTC ACAAATATAA CAAAGTTCAT TGTTTTAGGG TTTTTAAAAA  5100
  80    ACGTGTTGTA CCTAAGGCCA TACTTACTCT TCTATGCTAT CACTGCAAAG GGGTGATATG  5160
```

330

```
TATGTATTAT ATAAAAAAAA AAACCCTTAA TGCACTGTTA TCTCCTAAAT ATTTAGTAAA   5220
TTAATACTAT TTAATTTTTT TAAAGATTTG TCTGTGTAGA CACTAAAAGT ATTACACAAA   5280
ATCTGGACTG AAGGTGTCCT TTTTAACAAC AATTTAAAGT ACTTTTTATA TATGTTATGT   5340
AGTATATCCT TTCTAAACTG CCTAGTTTGT ATATTCCTAT AATTCCTATT TGTGAAGTGT   5400
ACCTGTTCTT GTCTCTTTTT TCAGTCATTT TCTGCACGCA TCCCCCTTTA TATGGTTATA   5460
GAGATGACTG TAGCTTTTCG TGCTCCACTG CGAGGTTTGT GCTCAGAGCC GCTGCACCCC   5520
AGCGAGGCCT GCTCCATGGA GTGCAGGACG AGCTACTGCT TTGGAGCGAG GGTTTCCTGC   5580
TTTTGAGTTG ACCTGACTTC CTTCTTGAAA TGACTGTTAA AACTAAAATA AATTACATTG   5640
CATTTATTTT ATATTCTTGG TTGAAATAAA ATTTAATTGA CTTTG
```

SEQ ID NO:78 PDO3 Protein sequence:
Protein Accession #:   BAA82980

```
1          11         21         31         41         51
|          |          |          |          |          |
VKSLLYQILD GIHYLHANWV LHRDLKPANI LVMGEGPERG RVKIADMGFA RLFNSPLKPL    60
ADLDPVVVTF WYRAPELLLG ARHYTKAIDI WAIGCIFAEL LTSEPIFHCR QEDIKTSNPF   120
HHDQLDRIFS VMGFPADKDW EDIRKMPEYP TLQKDFRRTT YANSSLIKYM EKHKVKPDSK   180
VFLLLQKLLT MDPTKRITSE QALQDPYFQE DPLPTLDVFA GCQIPYPKRE FLNEDDPEEK   240
GDKNQQQQQN QHQQPTAPPQ QAAAPPQAPP PQQNSTQTNG TAGGAGAGVG GTGAGLQHSQ   300
DSSLNQVPPN KKPRLGPSGA NSGGPVMPSD YQHSSSRLNY QSSVQGSSQS QSTLGYSSSS   360
QQSSQYHPSH QAHRY
```

                                   SEQ ID NO:79 PDO5 DNA SEQUENCE
Nucleic Acid Accession #:   XM_002922
Coding sequence: 1-2190 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGAATCCTT TCCAGAAAAA TGAGTCCAAG GAAACTCTTT TTTCACCTGT CTCCATTGAA    60
GAGGTACCAC CTCGACCACC TAGCCCTCCA AAGAAGCCAT CTCCGACAAT CTGTGGCTCC   120
AACTATCCAC TGAGCATTGC CTTCATTGTG GTGAATGAAT TCTGCGAGCG CTTTTCCTAT   180
TATGGAATGA AAGCTGTGCT GATCCTGTAT TTCCTGTATT TCCTGCACTG GAATGAAGAT   240
ACCTCCACAT CTATATACCA TGCCTTCAGC AGCCTCTGTT ATTTTACTCC CATCCTGGGA   300
GCAGCCATTG CTGACTCGTG GTTGGGAAAA TTCAAGACAA TCATCTATCT CTCCTTGGTG   360
TATGTGCTTG GCCATGTGAT CAAGTCCTTG GGTGCCTTAC CAATACTGGG AGGACAAGTG   420
GTACACACAG TCCTATCATT GATCGGCCTG AGTCTAATAG CTTTGGGGAC AGGAGGCATC   480
AAACCCTGTG TGGCAGCTTT TGGTGGAGAC CAGTTTGAAG AAAAACATGC AGAGGAACGG   540
ACTAGATACT TCTCAGTCTT CTACCTGTCC ATCAATGCAG GGAGCTTGAT TTCTACATTT   600
ATCACACCCA TGCTGAGAGG AGATGTGCAA TGTTTTGGAG AAGACTGCTA TGCATTGGCT   660
TTTGGAGTTC CAGGACTGCT CATGGTAATT GCACTTGTTG TGTTTGCAAT GGGAAGCAAA   720
ATATACAATA AACCACCCCC TGAAGGAAAC ATAGTGGCTC AAGTTTTCAA ATGTATCTGG   780
TTTGCTATTT CCAATCGTTT CAAGAACCGT TCTGGAGACA TTCCAAAGCG ACAGCACTGG   840
CTAGACTGGG CAGCTGAGAA ATATCCAAAG CAGCTCATTA TGGATGTAAA GGCACTGACC   900
AGGGTACTAT TCCTTTATAT CCCATTGCCC ATGTTCTGGG CTCTTTTGGA TCAGCAGGGT   960
TCACGATGGA CTTTGCAAGC CATCAGGATG AATAGGAATT TGGGGTTTTT TGTGCTTCAG   1020
CCGGACCAGA TGCAGGTTCT AAATCCCTTT CTGGTTCTTA TCTTCATCCC GTTGTTTGAC   1080
TTTGTCATTT ATCGTCTGGT CTCCAAGTGT GGAATTAACT TCTCATCACT TAGGAAAATG   1140
GCTGTTGGTA TGATCCTAGC GTGCCTGGCA TTTGCAGTTG CGGCAGCTGT AGAGATAAAA   1200
ATAAATGAAA TGGCCCCAGC CCAGTCAGGT CCCCAGGAGG TTTTCCTACA AGTCTTGAAT   1260
CTGGCAGATG ATGAGGTGAA GGTGACAGTG GTGGGAAATG AAAACAATTC TCTGTTGATA   1320
GAGTCCATCA AATCCTTTCA GAAAACACCA CACTATTCCA AACTGCACCT GAAAACAAAA   1380
AGCCAGGATT TTCACTTCCA CCTGAAATAT CACAATTTGT CTCTCTACAC TGAGCATTCT   1440
GTGCAGGAGA AGAACTGGTA CAGTCTTGTC ATTCGTGAAG ATGGGAACAG TATCTCCAGC   1500
ATGATGGTAA AGGATACAGA AAGCAAAACA ACCAATGGGA TGACAACCGT GAGGTTTGTT   1560
AACACTTTGC ATAAAGATGT CAACATCTCC CTGAGTACGA ATACCTCTCT CAATGTTGGT   1620
GAAGACTATG GTGTGTCTGC TTATAGAACT GTGCAAAGAG GAGAATACCC TGCAGTGCAC   1680
TGTAGAACAG AAGATAAGAA CTTTTCTCTG AATTTGGGTC TTCTAGACTT TGGTGCAGCA   1740
TATCTGTTTG TTATTACTAA TAACACCAAT CAGGGTCTTC AGGCCTGGAA GATTGAAGAC   1800
ATTCCAGCCA ACAAAATGTC CATTGCGTGG CAGCTACCAC AATATGCCCT GGTTACAGCT   1860
GGGGAGGTCA TGTTCTCTGT CACAGGTCTT GAGTTTTCTT ATTCTCAGGC TCCCTCTAGC   1920
ATGAAATCTG TGCTCCAGGC AGCTTGGCTA TTGACAATTG CAGTTGGGAA TATCATCGTG   1980
CTTGTTGTGG CACAGTTCAG TGGCCTGGTA CAGTGGGCCG AATTCATTTT GTTTTCCTGC   2040
CTCCTGCTGG TGATCTGCCT GATCTTCTCC ATCATGGGCT ACTACTATGT TCCTGTAAAG   2100
ACAGAGGATA TGCGGGGTCC AGCAGATAAG CACATTCCTC ACATCCAGGG GAACATGATC   2160
AAACTAGAGA CCAAGAAGAC AAAACTCTGA
```

SEQ ID NO:80 PDO5 Protein sequence:
Protein Accession #:   XP_002922

```
1          11         21         31         41         51
|          |          |          |          |          |
MNPFQKNESK ETLFSPVSIE EVPPRPPSPP KKPSPTICGS NYPLSIAFIV VNEFCERFSY    60
YGMKAVLILY FLYFLHWNED TSTSIYHAFS SLCYFTPILG AAIADSWLGK FKTIIYLSLV   120
YVLGHVIKSL GALPILGGQV VHTVLSLIGL SLIALGTGGI KPCVAAFGGD QFEEKHAEER   180
TRYFSVFYLS INAGSLISTF ITPMLRGDVQ CFGEDCYALA FGVPGLLMVI ALVVFAMGSK   240
IYNKPPPEGN IVAQVFKCIW FAISNRFKNR SGDIPKRQHW LDWAAEKYPK QLIMDVKALT   300
RVLFLYIPLP MFWALLDQQG SRWTLQAIRM NRNLGFFVLQ PDQMQVLNPF LVLIFIPLFD   360
FVIYRLVSKC GINFSSLRKM AVGMILACLA FAVAAVEIK INEMAPAQSG PQEVFLQVLN    420
LADDEVKVTV VGNENNSLLI ESIKSFQKTP HYSKLHLKTK SQDFHFHLKY HNLSLYTEHS   480
```

331

```
VQEKNWYSLV IREDGNSISS MMVKDTESKT TNGMTTVRFV NTLHKDVNIS LSTDTSLNVG     540
EDYGVSAYRT VQRGEYPAVH CRTEDKNFSL NLGLLDFGAA YLFVITNNTN QGLQAWKIED     600
IPANKMSIAW QLPQYALVTA GEVMFSVTGL EFSYSQAPSS MKSVLQAAWL LTIAVGNIIV     660
LVVAQFSGLV QWAEFILFSC LLLVICLIFS IMGYYYVPVK TEDMRGPADK HIPHIQGNMI     720
KLETKKTKL
```

SEQ ID NO:81 PDO6 DNA SEQUENCE
Nucleic Acid Accession #:   NM_020448
Coding sequence: 1-1221 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGGACGGAT CCCACAGCGC AGCCCTGAAG CTGCAGCAGC TGCCTCCCAC AAGTAGCTCC      60
AGCGCCGTAA GCGAGGCCTC CTTCTCCTAC AAGGAAAACC TGATTGGCGC CCTCTTGGCG     120
ATCTTCGGGC ACCTCGTGGT CAGCATTGCA CTTAACCTCC AGAAGTACTG CCACATCCGC     180
CTGGCAGGCT CCAAGGATCC CCGGGCCTAT TTCAAGACCA AGACATGGTG GCTGGGCCTG     240
TTCCTGATGC TTCTGGGCGA GCTGGGTGTG TTCGCCTCCT ACGCCTTCGC GCCGCTGTCA     300
CTCATCGTGC CCCTCAGCGC AGTTTCTGTG ATAGCTAGTG CCATCATAGG AATCATATTC     360
ATCAAGGAAA AGTGGAAACC GAAAGACTTT CTGAGGCGCT ACGTCTTGTC CTTTGTTGGC     420
TGCGGTTTGG CTGTCGTGGG TACCTACCTG CTGGTGACAT TCGCACCCAA CAGTCACGAG     480
AAGATGACAG GCGAGAATGT CACCAGGCAC CTCGTGAGCT GGCCTTTCCT TTTGTACATG     540
CTGGTGGAGA TCATTCTGTT CTGCTTGCTC CTCTACTTCT ACAAGGAGAA GAACGCCAAC     600
AACATTGTCG TGATTCTTCT CTTGGTGCGC TTACTTGGCT CCATGACAGT GGTGACAGTC     660
AAGGCCGTGG CTGGGATGCT TGTCTTGTCC ATTCAAGGGA ACCTGCAGCT TGACTACCCC     720
ATCTTCTACG TGATGTTCGT GTGCATGGTG GCAACCGCCG TCTATCAGGC TGCGTTTTTG     780
AGTCAAGCCT CACAGATGTA CGACTCCTCT TTGATTGCCA GTGTGGGCTA CATTCTGTCC     840
ACAACCATTG CTATCACAGC AGGTGCAATA TTTTACCTGG ACTTCATCGG GGAGGACGTG     900
CTGCACATCT GCATGTTTGC ACTGGGGTGC CTCATTGCAT TCTTGGGCGT CTTCTTAATC     960
ACGCGTAACA GGAAGAAGCC CATTCCATTT GAGCCCTATA TTTCCATGGA TGCCATGCCA    1020
GGTATGCAGA ACATGCACGA TAAAGGGATG ACTGTCCAGC CTGAACTTAA AGCTTCTTTT    1080
TCCTATGGGG CTCTGGAAAA CAATGACAAC ATTTCTGAGA TCTACGCTCC TGCCACCCTG    1140
CCAGTCATGC AAGAAGAGCA CGGCTCCAGA AGTGCCTCTG GGGTCCCCTA CCGAGTCCTA    1200
GAGCACACCA AGAAGGAATG A
```

SEQ ID NO:82 PDO6 Protein sequence
Protein Accession #:   NP_065181

```
1          11         21         31         41         51
|          |          |          |          |          |
MDGSHSAALK LQQLPPTSSS SAVSEASFSY KENLIGALLA IFGHLVVSIA LNLQKYCHIR      60
LAGSKDPRAY FKTKTWWLGL FLMLLGELGV FASYAFAPLS LIVPLSAVSV IASAIIGIIF     120
IKEKWKPKDF LRRYVLSFVG CGLAVVGTYL LVTFAPNSHE KMTGENVTRH LVSWPFLLYM     180
LVEIILFCLL LYFYKEKNAN NIVVILLLVA LLGSMTVVTV KAVAGMLVLS IQGNLQLDYP     240
IFYVMFVCMV ATAVYQAAFL SQASQMYDSS LIASVGYILS TTIAITAGAI FYLDFIGEDV     300
LHICMFALGC LIAFLGVFLI TRNRKKPIPF EPYISMDAMP GMQNMHDKGM TVQPELKASF     360
SYGALENNDN ISEIYAPATL PVMQEEHGSR SASGVPYRVL EHTKKE
```

SEQ ID NO:83 PDO8 DNA SEQUENCE
Nucleic Acid Accession #:   NM_032712
Coding sequence: 555-908 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
CACTCATTAA GAACAGAGGA GGCTGCCTGT TACTCCTGGT GTTGCATCCC TCCAGACACT      60
CTGCTGTTTC CTGCCTAGGC GTGGCTGCAG CCATGGCTAG GAAAGCGCTG CCACCCACCC     120
ACCTGGGCCA GAGCTGGTTC TGCTCCTGCT GCAGGGACAC TGAGCTGGCT ATCTCGGCGC     180
TTCGGGCAAG AACTGCAACA GGCTCTCCTG GGTCCTGCAG GTGTACAGCC GGGCCCCTGC     240
CTTGTGCCTC AGCTCTCGAG AGCTGCTGCT GCCGGGTGAC CTGATCCAAC CTGATAAGGT     300
GCCATCTTCA GCTACCACTG CAAGGCCCTG AGGGCAACGA CAGCACGGCA CTGCCCACCC     360
GGCTGCTGAT GGCCTGGTGC CAGCTGGGAG TCCTCCCGGC ACTTCGAGGC CACTGAGCCA     420
CCCTTCCAGC CCCAGCCCAC CATGGACAGG GGTATCCAGC TTCCTCCTCA ACCTCGTCCT     480
CTGCCCCTGA GCCAGTGACG CCCAAGGACA TGCCTGTTAC CCAGGTCCTG TACCAGCACT     540
AGCTGGTCAA GGGCATGACA GTGCTGGAGG CCGTCTTGGA GATCCAGGCC ATCACTGGCA     600
GCAGGCTGCT CTCCATGGTG CCAGGGCCCG CCAGGCCACC AGGCTCATGC TGGGACCCAA     660
CCCAGTGCAC AAGGACTTGG CTGCTGAGCC ACACACCCAG GAGAAGGTGG ATAAGTGGGC     720
TACCAAGGGC TTCCTGCAGG CTAGGGGAGG AGCCACCCCC GCTTCCCTAT TGTGACCAGG     780
CCTATGGGGA GGAGCTGTCC ATACGCCACC GTGAGACCTG GGCCTGGCTC TCAAGGACAG     840
ACACCGCCTG GCCTGGTGCT CCAGGGGTGA AGCAGGCCAG AATCCTGGGG GAGCTGCTCC     900
TGGTTTGAGC TGCATTCAGG AAGTGCGGGA CATGGTAGGG GAGGCAAAAA GCCTTGGGCA     960
CTACCCTCCC TGTGGAGCTG TTCGGTGTCC GTCGAGCTAG CCACACCCTG ACACCATGTT    1020
CAAGGGTACC GGAAGAGAAG GGTGTCTGCC CCCAACCTCC CCTGTGGGTG TCACTGGCCA    1080
GATGTCATGA GGGAAGCAGG CCTTGTGAGT GGACACTGAC CATGAGTCCC TGGGGGGAGT    1140
GATCCCCCAG GCATCGTGTG CCATGTTGCA CTTCTGCCCA GGCAGCAGGG TGGGTGGGTA    1200
CCATGGGTGC CCACCCCTCC ACCACATGGG GCCCCAAAGC ACTGCAGGCC AAGCAGGGCA    1260
ACCCCACACC CTTGACATAA AAGCATCTTG AAGCTTTTAA AAAAAAAAA AAAAAA
```

SEQ ID NO:84 PDO8 Protein sequence
Protein Accession #:   NP_116101

```
1          11         21         31         41         51
```

332

```
|             |             |             |             |             |             |
MTVLEAVLEI   QAITGSRLLS   MVPGPARPPG   SCWDPTQCTR   TWLLSHTPRR   RWISGLPRAS      60
CRLGEEPPPL   PYCDQAYGEE   LSIRHRETWA   WLSRTDTAWP   GAPGVKQARI   LGELLLV
```

SEQ ID NO:85 PDT1 DNA SEQUENCE

Nucleic Acid Accession #: NM_000693
Coding sequence: 53-1591 (underlined sequences correspond to start and stop codons)

```
1           11          21          31          41          51
|           |           |           |           |           |
AGCCGGTGCG  CCGCAGACTA  GGGCGCCTCG  GGCCAGGGAG  CGCGGAGGAG  CCATGGCCAC     60
CGCTAACGGG  GCCGTGGAAA  ACGGGCAGCC  GGACGGGAAG  CCGCCGGCCC  TGCCGCGCCC    120
CATCCGCAAC  CTGGAGGTCA  AGTTCACCAA  GATATTTATC  AACAATGAAT  GGCACGAATC    180
CAAGAGTGGG  AAAAAGTTTG  CTACATGTAA  CCCTTCAACT  CGGGAGCAAA  TATGTGAAGT    240
GGAAGAAGGA  GATAAGCCCG  ACGTGGACAA  GGCTGTGGAG  GCTGCACAGG  TTGCCTTCCA    300
GAGGGGCTCG  CCATGGCGCC  GGCTGGATGC  CCTGAGTCGT  GGGCGGCTGC  TGCACCAGCT    360
GGCTGACCTG  GTGGAGAGGG  ACCGCGCCAC  CTTGGCCGCC  CTGGAGACGA  TGGATACAGG    420
GAAGCCATTT  CTTCATGCTT  TTTTCATCGA  CCTGGAGGGC  TGTATTAGAA  CCCTCAGATA    480
CTTTGCAGGG  TGGGCAGACA  AAATCCAGGG  CAAGACCATC  CCCACAGATG  ACAACGTCGT    540
ATGCTTCACC  AGGCATGAGC  CCATTGGTGT  CTGTGGGGCC  ATCACTCCAT  GGAACTTCCC    600
CCTGCTGATG  CTGGTGTGGA  AGCTGGCACC  CGCCCTCTGC  TGTGGGAACA  CCATGGTCCT    660
GAAGCCTGCG  GAGCAGACAC  CTCTCACCGC  CCTTTATCTC  GGCTCTCTGA  TCAAAGAGGC    720
CGGGGTTCCCT  CCAGGAGTGG  TGAACATTGT  GCCAGGATTC  GGGCCCACAG  TGGGAGCAGC    780
AATTTCTTCT  CACCCTCAGA  TCAACAAGAT  CGCCTTCACC  GGCTCCACAG  AGGTTGGAAA    840
ACTGGTTAAA  GAAGCTGCGT  CCCGGAGCAA  TCTGAAGCGG  GTGACGCTGG  AGCTGGGGGG    900
GAAGAACCCC  TGCATCGTGT  GTGCGGACGC  TGACTTGGAC  TTGGCAGTGG  AGTGTGCCCA    960
TCAGGGAGTG  TTCTTCAACC  AAGGCCAGTG  TTGCACGGCA  GCCTCCAGGG  TGTTCGTGGA   1020
GGAGCAGGTC  TACTCTGAGT  TTGTCAGGCG  GAGCGTGGAG  TATGCCAAGA  AACGGCCCGT   1080
GGGGAGACCCC TTCGATGTCA  AAACAGAACA  GGGGCCTCAG  ATTGATCAAA  AGCAGTTCGA   1140
CAAAATCTTA  GAGCTGATCG  AGAGTGGGAA  GAAGGAAGGG  GCCAAGCTGG  AATGCGGGGG   1200
CTCAGCCATG  GAAGACAAGG  GGCTCTTCAT  CAAACCCACT  GTCTTCTCAG  AAGTCACAGA   1260
CAACATGCGG  ATTGCCAAAG  AGGAGATTTT  CGGGCCAGTG  CAACCAATAC  TGAAGTTCAA   1320
AAGTATCGAA  GAAGTGATAA  AAAGAGCGAA  TAGCACCGAC  TATGGACTCA  CAGCAGCCGT   1380
GTTCACAAAA  AATCTCGACA  AAGCCCTGAA  GTTGGCTTCT  GCCTTAGAGT  CTGGAACGGT   1440
CTGGATCAAC  TGCTACAACG  CCCTCTATGC  ACAGGCTCCA  TTTGGTGGCT  TTAAAATGTC   1500
AGGAAATGGC  AGAGAACTAG  GTGAATACGC  TTTGGCCGAA  TACACAGAAG  TGAAAACTGT   1560
CACCATCAAA  CTTGGCGACA  AGAACCCCTG  AAGGAAAGGC  GGGGCTCCTT  CCTCAAACAT   1620
CGGACGGCGG  AATGTGGCAG  ATGAAATGTG  CTGGAGGAAA  AAAATGACAT  TTCTGACCTT   1680
CCCGGGCACA  ATTCTTCTGG  AGGCTTTACA  TCTACTGGAG  TTGAATGATT  GCTGTTTTCC   1740
TCTCACTCTC  CTGTTTATTC  ACCAGACTGG  GGATGCCTAT  AGGTTGTCTG  TGAAATCGCA   1800
GTCCTGCCTG  GGGAGGGAGC  TGTTGGCCAT  TTCTGTGTTT  CCCTTTAAAC  CAGATCCTGG   1860
AGACAGTGAG  ATACTCAGGG  CGTTGTTAAC  AGGGAGTGGT  ATTTGAAGTG  TCCAGCAGTT   1920
GCTTGAAATG  CTTTGCCGAA  TCTGACTCCA  GTAAGAATGT  GGGAAAACCC  CCTGTGTGTT   1980
CTGCAAGCAG  GGCTCTTGCA  CCAGCGGTCT  CCTCAGGGTG  GACCTGCTTA  CAGAGCAAGC   2040
CACGCCTCTT  TCCGAGGTGA  AGGTGGGACC  ATTCCTTGGG  AAAGGATTCA  CAGTAAGGTT   2100
TTTTGGTTTT  TGTTTTTTGT  TTTCTTGTTT  TTAAAAAAAG  GATTTCACAG  TGAGAAAGTT   2160
TTGGTTAGTG  CATACCGTGG  AAGGGCGCCA  GGGTCTTTGT  GGATTGCATG  TTGACATTGA   2220
CCGTGAGATT  CGGCTTCAAA  CCAATACTGC  CTTTGGAATA  TGACAGAATC  AATAGCCCAG   2280
AGAGCTTAGT  CAAAGACGAT  ATCACGGTCT  ACCTTAACCA  AGGCACTTTC  TTAAGCAGAA   2340
AATATTGTTG  AGGTTACCTT  TGCTGCTAAA  GATCCAATCT  TCTAACGCCA  CAACAGCATA   2400
GCAAATCCTA  GGATAATTCA  CCTCCTCATT  TGACAAATCA  GAGCTGTAAT  TCACTTTAAC   2460
AAATTACGCA  TTTCTATCAC  GTTCACTAAC  AGCTTATGAT  AAGTCTGTGT  AGTCTTCCTT   2520
TTCTCCAGTT  CTGTTACCCA  ATTTAGATTA  GTAAAGCGTA  CACAACTGGA  AAGACTGCTG   2580
TAATACACA   GCCTTGTTAT  TTTTAAGTCC  TATTTTGATA  TTAATTTCTG  ATTAGTTAGT   2640
AAATAACACC  TGGATTCTAT  GGAGGACCTC  GGTCTTCATC  CAAGTGGCCT  GAGTATTTCA   2700
CTGGCAGGTT  GTGAATTTTT  CTTTTCCTCT  TTGGGAATCC  AAATGATGAT  GTGCAATTTC   2760
ATGTTTTAAC  TTGGGAAACT  GAAAGTGTTC  CCATATAGCT  TCAAAAACAA  AAACAAATGT   2820
GTTATCCGAC  GGATACTTTT  ATGGTTACTA  ACTAGTACTT  TCCTAATTGG  GAAAGTAGTG   2880
CTTAAGTTTG  CAAATTAAGT  TGGGGAGGGC  AATAATAAAA  TGAGGGCCCG  TAACAGAACC   2940
AGTGTGTGTA  TAACGAAAAC  CATGTATAAA  ATGGGCCTAT  CACCCTTGTC  AGAGATATAA   3000
ATTACCACAT  TTGGCTTCCC  TTCATCAGCT  AACACTTATC  ACTTATACTA  CCAATAACTT   3060
GTTAAATCAG  GATTTGGCTT  CATACACTGA  ATTTTCAGTA  TTTTATCTCA  AGTAGATATA   3120
GACACTAACC  TTGATAGTGA  TACGTTAGAG  GGTTCCTATT  CTTCCATTGT  ACGATAATGT   3180
CTTTAATATG  AAATGCTACA  TTATTTATAA  TTGGTAGAGT  TATTGTATCT  TTTTATAGTT   3240
GTAAGTACAC  AGAGGTGGTA  TATTTAAACT  TCTGTAATAT  ACTGTATTTA  GAAATGGAAA   3300
TATATATAGT  GTTAGGTTTC  ACTTCTTTTA  AGGTTTACCC  CTGTGGTGTG  GTTTAAAAAT   3360
CTATAGGCCT  GGGAATTCCG  ATCCTAGCTG  CAGATCGCAT  CCCACAATGC  GAGAATGATA   3420
AAATAAAATT  GGATATTTGA  GA
```

SEQ ID NO:86 PDT1 PROTEIN SEQUENCE

Protein Accession #:   NP_000684

```
1           11          21          31          41          51
|           |           |           |           |           |
MATANGAVEN  GQPDGKPPAL  PRPIRNLEVK  FTKIFINNEW  HESKSGKKFA  TCNPSTREQI     60
CEVEEGDKPD  VDKAVEAAQV  AFQRGSPWRR  LDALSRGRLL  HQLADLVERD  RATLAALETM    120
DTGKPFLHAF  FIDLEGCIRT  LRYFAGWADK  IQGKTIPTDD  NVVCFTRHEP  IGVCGAITPW    180
NFPLLMLVWK  LAPALCCGNT  MVLKPAEQTP  LTALYLGSLI  KEAGFPPGVV  NIVPGFGPTV    240
GAAISSHPQI  NKIAFTGSTE  VGKLVKEAAS  RSNLKRVTLE  LGGKNPCIVC  ADADLDLAVE    300
```

333

```
CAHQGVFFNQ GQCCTAASRV FVEEQVYSEF VRRSVEYAKK RPVGDPFDVK TEQGPQIDQK    360
QFDKILELIE SGKKEGAKLE CGGSAMEDKG LFIKPTVFSE VTDNMRIAKE EIFGPVQPIL    420
KFKSIEEVIK RANSTDYGLT AAVFTKNLDK ALKLASALES GTVWINCYNA LYAQAPFGGF    480
KMSGNGRELG EYALAEYTEV KTVTIKLGDK NP
```

SEQ ID NO:87 PDV3 DNA SEQUENCE

Nucleic Acid Accession #:   NM_032642
Coding sequence: 184-1263 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
GACCATTAGC AGGCACCCAG GCCTGTCTTT GGCTCGGAAA CGGTGGCCCC CAATGTAGCC     60
TAGTTTGAAC CTAGGAACTG CAGGACCAGA GAGATTCCAC TGGAGCCTGA TGGACGGGTG    120
ACAGAGGGAA CCCTACTCTG GAAACTGTCA GTCCCAGGGC ACTGGGGAGG GCTGAGGCCG    180
ACCATGCCCA GCCTGCTGCT GCTGTTCACG GCTGCTCTGC TGTCCAGCTG GGCTCAGCTT    240
CTGACAGACG CCAACTCCTG GTGGTCATTA GCTTTGAACC CGGTGCAGAG ACCCGAGATG    300
TTTATCATCG GTGCCCAGCC CGTGTGCAGT CAGCTTCCCG GGCTCTCCCC TGGCCAGAGG    360
AAGCTGTGCC AATTGTACCA GGAGCACATG GCCTACATAG GGGAGGGAGC CAAGACTGGC    420
ATCAAGGAAT GCCAGCACCA GTTCCGGCAG CGGCGGTGGA ATTGCAGCAC AGCGGACAAC    480
GCATCTGTCT TTGGGAGAGT CATGCAGATA GGCAGCCGAG AGACCGCCTT CACCCACGCG    540
GTGAGCGCCG CGGGCGTGGT CAACGCCATC AGCCGGGCCT GCCGCGAGGG CGAGCTCTCC    600
ACCTGCGGCT GCAGCCGGAC GGCGCGGCCC AAGGACCTGC CCCGGGACTG GCTGTGGGGC    660
GGCTGTGGGG ACAACGTGGA GTACGGCTAC CGCTTCGCCA AGGAGTTTGT GGATGCCCGG    720
GAGCGAGAGA AGAACTTTGC CAAAGGATCA GAGGAGCAGG GCCGGGTGCT CATGAACCTG    780
CAAAACAACG AGGCCGGTCG CAGGGCTGTG TATAAGATGG CAGACGTAGC CTGCAAATGC    840
CACGGCGTCT CGGGGTCCTG CAGCCTCAAG ACCTGCTGGC TGCAGCTGGC CGAGTTCCGC    900
AAGGTCGGGG ACCGGCTGAA GGAGAAGTAC GACAGCGCGG CCGCCATGCG CGTCACCCGC    960
AAGGGCCGGC TGGAGCTGGT CAACAGCCGC TTCACCCAGC CCACCCCGGA GGACCTGGTC   1020
TATGTGGACC CCAGCCCCGA CTACTGCCTG CGCAACGAGA GCACGGGCTC CCTGGGCACG   1080
CAGGGCCGCC TCTGCAACAA GACCTCGGAG GGCATGGATG GCTGTGAGCT CATGTGCTGC   1140
GGGCGTGCCT ACAACCAGTT CAAGAGCGTG CAGGTGGAGC GCTGCCACTG CAAGTTCCAC   1200
TGGTGCTGCT TCGTCAGGTG TAAGAAGTGC ACGGAGATCG TGGACCAGTA CATCTGTAAA   1260
TAGCCCGGAG GGCCTGCTCC CGGCCCCCCC TGCACTCTGC CTCACAAAGG TCTATATTAT   1320
ATAAATCTAT ATAAATCTAT TTTATATTTG TATAAGTAAA TGGGTGGGTG CTATACAATG   1380
GAAAGATGAA AATGGAAAGG AAGAGCTTAT TTAAGACACG CTGGAGATCT CTGAGGAGTG   1440
GACTTTGCTG GTTCTCTCCT CTTGGTGGGT GGGAGACAGG GCTTTTTCTC TCCCTCTGGC   1500
GAGGACTCTC AGGATGTAGG GACTTGGAAA TATTTACTGT CTGTCCACCA CGGCCTGGAG   1560
GAGGGAGGTT GTGGTTGGAT GGAGGAGATG ATCTTGTCTG GAAGTCTAGA GTCTTTGTTG   1620
GTTAGAGGAC TGCCTGTGAT CCTGGCCACT AGGCCAAGAG GCCCTATGAA GGTGGCGGGA   1680
ACTCAGCTTC AACCTCGATG TCTTCAGGGT CTTGTCCAGA ATGTAGATGG GTTCCGTAAG   1740
AGGCCTGGTG CTCTCTTACT CTTTCATCCA CGTGCACTTG TGCGGCATCT GCAGTTTACA   1800
GGAACGGCTC CTTCCCTAAA ATGAGAAGTC CAAGGTCATC TCTGGCCCAG TGACCACAGA   1860
GAGATCTGCA CCTCCCGGAC TTCAGGCCTG CCTTTCCAGC GAGAATTCTT CATCCTCCAC   1920
GGTTCACTAG CTCCTACCTG AAGAGGAAAG GGGGCCATTT GACCTGACAT GTCAGGAAAG   1980
CCCTAAACTG AATGTTTGCG CCTGGGCTGC AGAAGCCAGG GTGCATGACC AGGCTGCGTG   2040
GACGTTATAC TGTCTTCCCC CACCCCCGGG GAGGGGAAGC TTGAGCTGCT GCTGTCACTC   2100
CTCCACCGAG GGAGGCCTCA CAAACCACAG GACGCTGCAA CGGGTCAGGC TGGCGGGCCC   2160
GGCGTGCTCA TCATCTCTGC CCCAGGTGTA CGGTTTCTCT CTGACATTAA ATGCCCTTCA   2220
TGGAAAAAAA AAAAAGAAAA AAAAAAAAAA AA
```

SEQ ID NO:88 PDV3 Protein sequence
Protein Accession #:   NP_116031

```
1          11         21         31         41         51
MPSLLLLFTA ALLSSWAQLL TDANSWWSLA LNPVQRPEMF IIGAQPVCSQ LPGLSPGQRK     60
LCQLYQEHMA YIGEGAKTGI KECQHQFRQR RWNCSTADNA SVFGRVMQIG SRETAFTHAV    120
SAAGVVNAIS RACREGELST CGCSRTARPK DLPRDWLWGG CGDNVEYGYR FAKEFVDARE    180
REKNFAKGSE EQGRVLMNLQ NNEAGRRAVY KMADVACKCH GVSGSCSLKT CWLQLAEFRK    240
VGDRLKEKYD SAAAMRVTRK GRLELVNSRF TQPTPEDLVY VDPSPDYCLR NESTGSLGTQ    300
GRLCNKTSEG MDGCELMCCG RGYNQFKSVQ VERCHCKFHW CCFVRCKKCT EIVDQYICK
```

SEQ ID NO:89 PDT9 DNA SEQUENCE

Nucleic Acid Accession #:   NM_033280
Coding sequence: 58-636 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
GGCAGCCGTC TGTGCCACCC AGAGCCGGCG GGCCGCTAGG TCCCCGGAGA CCCTGCTATG     60
GTGCGTGCGG GCGCCGTGGG GGCTCATCTC CCCGCGTCCG GCTTGGATAT CTTCGGGGAC    120
CTGAAGAAGA TGAACAAGCG CCAGCTCTAT TACCAGGTTT TAAACTTCGC CATGATCGTG    180
TCTTCTGCAC TCATGATATG GAAAGGCTTG ATCGTGCTCA CAGGCAGTGA GAGCCCCATC    240
GTGGTGGTGC TGAGTGGCAG TATGGAGCCG GCCTTTCACA GAGGAGACCT CCTGTTCCTC    300
ACAAATTTCC GGGAAGACCC AATCAGAGCT GGTGAAATAG TTGTTTTTAA AGTTGAAGGA    360
CGAGACATTC CAATAGTTCA CAGAGTAATC AAAGTTCATG AAAAAGATAA TGGAGACATC    420
AAATTTCTGA CTAAAGGAGA TAATAATGAA GTTGATGATA GAGGCTTGTA CAAAGAAGGC    480
CAGAACTGGC TGGAAAAGAA GGACGTGGTG GGAAGAGCAA GAGGGTTTTT ACCATATGTT    540
GGTATGGTCA CCATAATAAT GAATGACTAT CCAAAATTCA AGTATGCTCT TTTGGCTGTA    600
ATGGGTGCAT ATGTGTTACT AAAACGTGAA TCCTAAAATG AGAAGCAGTT CCTGGGACCA    660
GATTGAAATG AATTCTGTTG AAAAAGAGAA AAACTAATAT ATTTGAGATG TTCCATTTTC    720
```

334

```
TGTATAAAAG GGAACAGTGT GGAGATGTTT TTGTCTTGTC CAAATAAAAG ATTCACCAGT   780
AAAAAAAAAA AAAA
```

SEQ ID NO:90 PDT9 Protein sequence
Protein Accession #:   NP_150596

```
1          11         21         31         41         51
|          |          |          |          |          |
MVRAGAVGAH LPASGLDIFG DLKKMNKRQL YYQVLNFAMI VSSALMIWKG LIVLTGSESP    60
IVVVLSGSME PAFHRGDLLF LTNFREDPIR AGEIVVFKVE GRDIPIVHRV IKVHEKDNGD   120
IKFLTKGDNN EVDDRGLYKE GQNWLEKKDV VGRARGFLPY VGMVTIIMND YPKFKYALLA   180
VMGAYVLLKR ES
```

SEQ ID NO:91 PDV5 DNA SEQUENCE
Nucleic Acid Accession #:   NM_016590
Coding sequence: 691-975 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
GATTACTCAC ACAGTCTTGA AGATGCAATG TCAGCTATTT AGGACAGAAA CATCCAAGGC    60
CGTGTCAGAA CTCAATTACG ACTACATATG CATTAAGGCA GGAACTGGCA GGCCTCAGGG   120
TACGCCAACT ATAGGACTCG TGCTTCTCGT ACGCTGGGCT ATAATCTATG AAACTGAGCT   180
CCAGAGCCAG CCAATCACTT AGCTCCTCAT AACAAGTCTA ACTGGCTCTG GAAAGCTGAA   240
AGGGCTGCAC TGGAACAACA CAGATGAGAT ATTCTACACA TTAATCTACT TATCTGGAAT   300
CACTTTGCCT CTAAAGGCCA GAGAAAAATC ACAGCTTCCT TGTCGGAGGG GAAAAGGACA   360
GGTGATCTGG GGAAAACGCA GCTACACCTG GAGCAAGGTC TCTTCCCGGC TTGGCAATCT   420
CAGCTGTGCC GGCGCTACGG GACCCGAGCC GTCCCAGAAA CCAAAGGGCA GGCACGGCAG   480
CAAACGCCTG AGTGCTGCTG CCTTCGGTGA CTATATGAGA ATGGAAACTT CTAAGGAAGC   540
CAGGTTGTTA GAATTGTTAC CCCCTTTACT CAGAGATAAC ATAGATTATC CAGGCTGAGA   600
TGGAAAACAA GCCCTTTATT GAATTTTCAA CACAGACTCC CTGCTTCTCA TCTCCTTAAT   660
AAAATTTCAT TAAAATCCCC TTGAACTCCC ATGTTCAAAT CTCCATTTGT TGACAGACAA   720
AGCCAACAAT ACTCTAAACT GAGGCCTGCA AGTCATTTCA TTTGTATTTT TGTCCAGAAA   780
TTTCCCATAG GAAGACTTCA CCTCCTACAA CTCCGAAGAA AACCCTTACT GTCCAAGACC   840
GTCACCAGCA ACCATCCGCA GTCATTCAAG TGGAAGCTTT CACAGCTTTT GTACATTCTC   900
TGTGTCAATA TACAACTGAG TTACAGACTG TCCCCTGACC CCCTGACCCT TACAAACACT   960
AAAAGTTTTG TTTGACTCAA CTTCAAGCTG CTCATCTGTT AGTAAGTGAT GTTCACTCCA  1020
GAACACATTC ATGATGAGAA CTTTCTAAAA GACCAGCACT GCTCTTCCCC TCCTATAATC  1080
ATAATAATCA TGATAACCTG AAACATGTTA CTGGGACTCG ACATTTTTCT GGGGATTGAA  1140
ATCTTTAGTC CTTGGAGCTG TCACATAGCA GGGGCAACCT CACACTGAAA CAAAGGAAGT  1200
GATGTCCCAT TATTATCCAC CCTGAGCCAC CATAATATGC TGTTTACATT TATTTTCTTC  1260
AGCCTGTGCA AAACAAAGCA ATGGAAAAGG AAACTAAAAA ATATACATAC TAGTACCATT  1320
ATCTTCTTTT GCCTAAAATT ACTAATGCAC CACGTCAGTC TGCTTCCTTC AGGCATCATT  1380
CTCAATTCAT CAGGACTTGT ATTAGCAGGT TCTGGCTAGA GAGACTATCT CCTGTCATCA  1440
CGATCAATTA ATGTTTTCTG GTGATCACAT CAGGCCCTAT CTAAGAAGCT CATGGTATAC  1500
AAGGGTCACC CAAATGACTG AGTGCAGTCC TTGCTCATAT TTCCTTCATC TTAACCCCGC  1560
AAACAAGAAT TAAGATGATC CCAATAAAAG AAAAAATTGCT CAGGAAACTG AACCTTTTTC  1620
TGAACCAAGC ACTGTCAGCA AATCTCAGGT ATTAGAGCAA CTATGGTTGA TTGAAAAGTG  1680
TCTCAAAATC TGGGCCAAGA ATGATTGCTA GGTCCATAAG CTAATTTGTC TGGCCTTGCC  1740
ATTTACGTAA GCCAAAGAAA GTCACTCATG AGTAAACTAT AGAAAACGTT CAGACCCATC  1800
CTGTTAGTAT GTCAAATCAA CTAAGACTGG CAGGGTATTA ACTCCATTCC AGGTGACATG  1860
GATAAAGAGC CCCATTATTT TCACAGTGCC AGCCTCTACC TAAGGAAACC CTAGACCTTG  1920
GAACCAGTTT CCTGGTAGGG AACTGCTGAC AGTTTCAATG CTGACAGTTG GAGCCAATGC  1980
CTCATAGTGT AAACTGAAAG AAAAATAGTT GCTTTTTAAA ATGTCAGCAA GAAGGCCTGC  2040
CTCATCTTAA CAAAGCAAAA AAAAATGCTT TAATTCAAAT TAAAAATCAT GATACTAAAA  2100
AAAAAAAA
```

SEQ ID NO:92 PDV5 Protein sequence
Protein Accession #:   NP_057674

```
1          11         21         31         41         51
|          |          |          |          |          |
MQCQLFRTET SKAVSELNYD YICIKAGTGR PQGTPTIGLV LLVRWAIIYE TELQSQPIT
```

SEQ ID NO:93 PEE6 DNA SEQUENCE
Nucleic Acid Accession #:   NM_002606
Coding sequence: 61-1842 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
CGCGGCGGCT GGCGTCGGGA AAGTACAGTA AAAAGTCCGA GTGCAGCCGC CGGGCGCAGG    60
ATGGGATCCG GCTCCTCCAG CTACCGGCCC AAGGCCATCT ACCTGGACAT CGATGGACGC   120
ATTCAGAAGG TAATCTTCAG CAAGTACTGC AACTCCAGCG ACATCATGGA CCTGTTCTGC   180
ATCGCCACCG GCCTGCCTCG GAACACGACC ATCTCCCTGC TGACCACCGA CGACGCCATG   240
GTCTCCATCG ACCCCACCAT GCCCGCGAAT TCAGAACGCA CTCCGTACAA AGTGAGACCT   300
GTGGCCATCA AGCAACTCTC CGCTGGTGTC GAGGACAAGA GAACCACAAG CCGTGGCCAG   360
TCTGCTGAGA GACCACTGAG GGACAGACGG GTTGTGGGCC TGGAGCAGCC CCGGAGGGAA   420
GGAGCATTTG AAAGTGGACA GGTAGAGCCC AGGCCCAGAG AGCCCCAGGG CTGCTACCAG   480
GAAGGCCAGC GCATCCCTCC AGAGAGAGAA GAATTAATCC AGAGCGTGCT GGCGCAGGTT   540
GCAGAGCAGT TCTCAAGAGC ATTCAAAATC AATGAACTGA AAGCTGAAGT TGCAAATCAC   600
TTGGCTGTCC TAGAGAAACG CGTTGGAATT G GAAGGACTAA AAGTGGTGGA GATTGAGAAA   660
```

```
TGCAAGAGTG ACATTAAGAA GATGAGGGAG GAGCTGGCGG CCAGAAGCAG CAGGACCAAC    720
TGCCCCTGTA AGTACAGTTT TTTGGATAAC CACAAGAAGT TGACTCCTCG ACGCGATGTT    780
CCCACTTACC CCAAGTACCT GCTCTCTCCA GAGACCATCG AGGGCCTGCG GAAGCCGACC    840
TTTGACGTCT GGCTTTGGGA GCCCAATGAG ATGCTGAGCT GCCTGGAGCA CATGTACCAC    900
GACCTCGGGC TGGTCAGGGA CTTCAGCATC AACCCTGTCA CCCTCAGGAG GTGGCTGTTC    960
TGTGTCCACG ACAACTACAG AAACAACCCC TTCCACAACT TCCGGCACTG CTTCTGCGTG   1020
GCCCAGATGA TGTACAGCAT GGTCTGGCTC TGCAGTCTCC AGGAGAAGTT CTCACAAACG   1080
GATATCCTGA TCCTAATGAC AGCGGCCATC TGCCACGATC TGGACCATCC CGGCTACAAC   1140
AACACGTACC AGATCAATGC CCGCACAGAG CTGGCGGTCC GCTACAATGA CATCTCACCG   1200
CTGGAGAACC ACCACTGCGC CGTGGCCTTC CAGATCCTCG CCGAGCCTGA GTGCAACATC   1260
TTCTCCAACA TCCCACCTGA TGGGTTCAAG CAGATCCGAC AGGGAATGAT CACATTAATC   1320
TTGGCCACTG ACATGGCAAG ACATGCAGAA ATTATGGATT CTTTCAAAGA GAAAATGGAG   1380
AATTTTGACT ACAGCAACGA GGAGCACATG ACCCTGCTGA AGATGATTTT GATAAAATGC   1440
TGTGATATCT CTAACGAGGT CCGTCCAATG GAAGTCGCAG AGCCTTGGGT GGACTGTTTA   1500
TTAGAGGAAT ATTTTATGCA GAGCGACCGT GAGAAGTCAG AAGGCCTTCC TGTGGCACCG   1560
TTCATGGACC GAGACAAAGT GACCAAGGCC ACAGCCCAGA TTGGGTTCAT CAAGTTTGTC   1620
CTGATCCCAA TGTTTGAAAC AGTGACCAAG CTCTTCCCCA TGGTTGAGGA GATCATGCTG   1680
CAGCCACTTT GGGAATCCCG AGATCGCTAC GAGGAGCTGA AGCGGATAGA TGACGCCATG   1740
AAAGAGTTAC AGAAGAAGAC TGACAGCTTG ACGTCTGGGG CCACCGAGAA GTCCAGAGAG   1800
AGAAGCAGAG ATGTGAAAAA CAGTGAAGGA GACTGTGCCT GAGGAAAGCG GGGGGCGTGG   1860
CTGCAGTTCT GGACGGGCTG GCCGAGCTGC GCGGGATCCT TGTGCAGGGA AGAGCTGCCC   1920
TGGGCACCTG GCACCACAAG ACCATGTTTT CTAAGAACCA TTTTGTTCAC TGATACAAAA   1980
AAAAAAAAA A
```

SEQ ID NO:94 PEE6 Protein sequence
Protein Accession #:   NP_002597

```
1          11         21         31         41         51
|          |          |          |          |          |
MGSGSSSYRP KAIYLDIDGR IQKVIFSKYC NSSDIMDLFC IATGLPRNTT ISLLTTDDAM    60
VSIDPTMPAN SERTPYKVRP VAIKQLSAGV EDKRTTSRGQ SAERPLRDRR VVGLEQPRRE   120
GAFESGQVEP RPREPQGCYQ EGQRIPPERE ELIQSVLAQV AEQFSRAFKI NELKAEVANH   180
LAVLEKRVEL EGLKVVEIEK CKSDIKKMRE ELAARSSRTN CPCKYSFLDN HKKLTPRRDV   240
PTYPKYLLSP ETIEALRKPT FDVWLWEPNE MLSCLEHMYH DLGLVRDFSI NPVTLRRWLF   300
CVHDNYRNNP FHNFRHCFCV AQMMYSMVWL CSLQEKFSQT DILILMTAAI CHDLDHPGYN   360
NTYQINARTE LAVRYNDISP LENHHCAVAF QILAEPECNI FSNIPPDGFK QIRQGMITLI   420
LATDMARHAE IMDSFKEKME NFDYSNEEHM TLLKMILIKC CDISNEVRPM EVAEPWVDCL   480
LEEYFMQSDR EKSEGLPVAP FMDRDKVTKA TAQIGFIKFV LIPMFETVTK LFPMVEEIML   540
QPLWESRDRY EELKRIDDAM KELQKKTDSL TSGATEKSRE RSRDVKNSEG DCA
```

SEQ ID NO:95 PEG4 DNA SEQUENCE

Nucleic Acid Accession #:   none
Coding sequence: 41-559 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
CAGTCACAGG CGAGAGCCYT GGGATGCACC GGCCAGAGGC ATGCTGCTGC TGCTCACGCT    60
TGCCCTCCTG GGGGGCCCCA CCTGGGCAGG GAAGATGTAT GGCCCTGGAG GAGGCAAGTA   120
TTTCAGCACC ACTGAAGACT ACGACCATGA AATCACAGGG CTGCGGGTGT CTGTAGGTCT   180
TCTCCTGGTG AAAAGTGTCC AGGTGAAACT TGGAGACTCC TGGGACGTGA AACTGGGAGC   240
CTTAGGTGGG AATACCCAGG AAGTCACCCT GCAGCCAGGC GAATACATCA CAAAAGTCTT   300
TGTCGCCTTC CAAGCTTTCC TCCGGGGTAT GGTCATGTAC ACCAGCAAGG ACCGCTATTT   360
CTATTTTGGG AAGCTTGATG GCCAGATCTC CTCTGCCTAC CCCAGCCAAG AGGGGCAGGT   420
GCTGGTGGGC ATCTATGGCC AGTATCAACT CCTTGGCATC AAGAGCATTG GCTTTGAATG   480
GAATTATCCA CTAGAGGAGC CGACCACTGA GCCACCAGTT AATCTCACAT ACTCAGCAAA   540
CTCACCCGTG GGTCGCTAGG GTGGGGTATG GGGCCATCCG AGCTGAGGCC ATCTGTGTGG   600
TGGTGGCTGA TGGTACTGGA GTAACTGAGT CGGGACGCTG AATCTGAATC CACCAATAAA   660
TAAAGCTTCT GCAGAATCAG TGAAAAAAAA A
```

SEQ ID NO:96 PEG4 Protein sequence
Protein Accession #:   FGENESH predicted

```
1          11         21         31         41         51
|          |          |          |          |          |
MLLLLTLALL GGPTWAGKMY GPGGGKYFST TEDYDHEITG LRVSVGLLLV KSVQVKLGDS    60
WDVKLGALGG NTQEVTLQPG EYITKVFVAF QAFLRGMVMY TSKDRYFYFG KLDGQISSAY   120
PSQEGQVLVG IYGQYQLLGI KSIGFEWNYP LEEPTTEPPV NLTYSANSPV GR
```

SEQ ID NO:97 PEL9 DNA SEQUENCE

Nucleic Acid Accession #:   NM_006953
Coding sequence: 33-896(underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
CCGTTCCGCG CTCTGGCGGC TCCTCCCGGG CGATGCCTCC GCTCTGGGCC CTGCTGGCCC    60
TCGGCTGCCT GCGGTTCGGC TCGGCTGTGA ACCTGCAGCC CCAACTGGCC AGTGTGACTT   120
TCGCCACCAA CAACCCCACA CTTACCACTG TGGCCTTGGA AAAGCCTCTC TGCATGTTTG   180
ACAGCAAAGA GGCCCTCACT GGCACCCACG AGGTCTACCT GTATGTCCTG GTCGACTCAG   240
CCATTTCCAG GAATGCCTCA GTGCAAGACA GCACCAACAC CCCACTGGGC TCAACGTTCC   300
```

336

```
TACAAACAGA GGGTGGGAGG ACAGGTCCCT ACAAAGCTGT GGCCTTTGAC CTGATCCCCT   360
GCAGTGACCT GCCCAGCCTG GATGCCATTG GGGATGTGTC CAAGGCCTCA CAGATCCTGA   420
ATGCCTACCT GGTCAGGGTG GGTGCCAACG GGACCTGCCT GTGGGATCCC AACTTCCAGG   480
GCCTCTGTAA CGCACCCCTG TCGGCAGCCA CGGAGTACAG GTTCAAGTAT GTCCTGGTCA   540
ATATGTCCAC GGGCTTGGTA GAGGACCAGA CCCTGTGGTC GGACCCCATC CGCACCAACC   600
AGCTCACCCC ATACTCGACG ATCGACACGT GGCCAGGCCG GCGGGAGCGGA GGCATGATCG   660
TCATCACTTC CATCCTGGGC TCCCTGCCCT TCTTTCTACT TGTGGGTTTT GCTGGCGCCA   720
TTGCCCTCAG CCTCGTGGAC ATGGGGAGTT CTGATGGGGA AACGACTCAC GACTCCCAAA   780
TCACTCAGGA GGCTGTTCCC AAGTCGCTGG GGGCCTCGGA GTCTTCCTAC ACGTCCGTGA   840
ACCGGGGGCC GCCACTGGAC AGGGCTGAGG TGTATTCCAG CAAGCTCCAA GACTGAGCCC   900
AGCACCACCC CTGGGCAGCA GCATCCTCCT CTCTGGCCTT GCCCCAGGCC CTGCAGCGGT   960
GGTTGTCACA CCCTGACTTC AGGGAAGGTG AAACAGGGCT TGTCCCTCCA ACTGCAGGAA  1020
AACCCTTAAT AAAATCTTCT GATGAGTTCT AAAAAAAAA
```

SEQ ID NO:98 <u>PEL9 Protein sequence</u>
Protein Accession #:   NP_008884

```
1         11        21        31        41        51
|         |         |         |         |         |
MPPLWALLAL GCLRFGSAVN LQPQLASVTF ATNNPTLTTV ALEKPLCMFD SKEALTGTHE    60
VYLYVLVDSA ISRNASVQDS TNTPLGSTFL QTEGGRTGPY KAVAFDLIPC SDLPSLDAIG   120
DVSKASQILN AYLVRVGANG TCLWDPNFQG LCNAPLSAAT EYRFKYVLVN MSTGLVEDQT   180
LWSDPIRTNQ LTPYSTIDTW PGRRSGGMIV ITSILGSLPF FLLVGFAGAI ALSLVDMGSS   240
DGETTHDSQI TQEAVPKSLG ASESSYTSVN RGPPLDRAEV YSSKLQD
```

SEQ ID NO:99 PEN1 DNA SEQUENCE
Nucleic Acid Accession #:   NM_012391
Coding sequence: 416-1423 (underlined sequences correspond to start and stop codons)

```
1         11        21        31        41        51
|         |         |         |         |         |
GTCTGACTTC CTCCCAGCAC ATTCCTGCAC TCTGCCGTGT CCACACTGCC CCACAGACCC    60
AGTCCTCCAA GCCTGCTGCC AGCTCCCTGC AAGCCCCTCA GGTTGGGCCT TGCCACGGTG   120
CCAGCAGGCA GCCCTGGGCT GGGGGTAGGG GACTCCCTAC AGGCACGCAG CCCTGAGACC   180
TCAGAGGGCC ACCCCTTGAG GGTGGCCAGG CCCCCAGTGG CCAACCTGAG TGCTGCCTCT   240
GCCACCAGCC CTGCTGGCCC CTGGTTCCGC TGGCCCCCCA GATGCCTGGC TGAGACACGC   300
CAGTGGCCTC AGCTGCCCAC ACCTCTTCCC GGCCCCTGAA GTTGGCACTG CAGCAGACAG   360
CTCCCTGGGC ACCAGGCAGC TAACAGACAC AGCCGCCAGC CCAAACAGCA GCGGCATGGG   420
CAGCGCCAGC CCGGGTCTGA GCAGCGTATC CCCCAGCCAC CTCCTGCTGC CCCCCGACAC   480
GGTGTCGCGG ACAGGCTTGG AGAAGGCGGC AGCGGGGGCA GTGGGTCTCG AGAGACGGGA   540
CTGGAGTCCC AGTCCACCCG CCACGCCCGA GCAGGGCCTG TCCGCCTTCT ACCTCTCCTA   600
CTTTGACATG CTGTACCCTG AGGACAGCAG CTGGGCAGCC AAGGCCCCTG GGGCCAGCAG   660
TCGGGAGGAG CCACCTGAGG AGCCTGAGCA GTGCCCGGTC ATTGACAGCC AAGCCCCAGC   720
GGGCAGCCTG GACTTGGTGC CCGGCGGGCT GACCTTGGAG GAGCACTCGC TGGAGCAGGT   780
GCAGTCCATG GTGGTGGGCG AAGTGCTCAA GGACATCGAG ACGGCCTGCA AGCTGCTCAA   840
CATCACCGCA GATCCCATGG ACTGGAGCCC CAGCAATGTG CAGAAGTGGC TCCTGTGGAC   900
AGAGCACCAA TACCGGCTGC CCCCCATGGG CAAGGCCTTC CAGGAGCTGG CGGGCAAGGA   960
GCTGTGCGCC ATGTCGGAGG AGCAGTTCCG CCAGCGCTCT CCCCTGGGTG GGGATGTGCT  1020
GCACGCCCAC CTGGACATCT GGAAGTCAGC GGCCTGGATG AAAGAGCGGA CTTCACCTGG  1080
GGCGATTCAC TACTGTGCCT CGACCAGTGA GGAGAGCTGG ACCGACAGCG AGGTGGACTC  1140
ATCATGCTCC GGGCAGCCCA TCCACCTGTG GCAGTTCCTC AAGGAGTTGC TACTCAAGCC  1200
CCACAGCTAT GGCCGCTTCA TTAGGTGGCT CAACAAGGAG AAGGGCATCT TCAAAATTGA  1260
GGACTCAGCC CAGGTGGCCC GGCTGTGGGG CATCCGCAAG AACCGTCCCG CCATGAACTA  1320
CGACAAGCTG AGCCGCTCCA TCCGCCAGTA TTACAAGAAG GGCATCATCC GGAAGCCAGA  1380
CATCTCCCAG CGCCTCGTCT ACCAGTTCGT GCACCCCATC TGAGTGCCTG GCCCCAGGGCC  1440
TGAAACCCGC CCTCAGGGGC CTCTCTCCTG CCTGCCCTGC CTCAGCCAGG CCCTGAGATG  1500
GGGGAAAACG GGCAGTCTGC TCTGCTGCTC TGACCTTCCA GAGCCCAAGG TCAGGGAGGG  1560
GCAACCAACT GCCCCAGGGG GATATGGGTC CTCTGGGGCG TTCGGGACCA TGGGGCAGGG  1620
GTGCTTCCTC CTCAGGCCCA GCTGCTCCCC TGGAGGACAG AGGGAGACAG GGCTGCTCCC  1680
CAACACCTGC CTCTGACCCC AGCATTTCCA GAGCAGAGCC TACAGAAGGG CAGTGACTCG  1740
ACAAAGGCCA CAGGCAGTCC AGGCCTCTCT CTGCTCCATC CCCCTGCCTC CCATTCTGCA  1800
CCACACCTGG CATGGTGCAG GGAGACATCT GCACCCCTGA GTTGGGCAGC CAGGAGTGCC  1860
CCCGGGAATG GATAATAAAG ATACTAGAGA ACTG
```

SEQ ID NO:100 <u>PEN1 Protein sequence</u>
Protein Accession #:   NP_036523

```
1         11        21        31        41        51
|         |         |         |         |         |
MGSASPGLSS VSPSHLLLPP DTVSRTGLEK AAAGAVGLER RDWSPSPPAT PEQGLSAFYL    60
SYFDMLYPED SSWAAKAPGA SSREEPPEEP EQCPVIDSQA PAGSLDLVPG GLTLEEHSLE   120
QVQSMVVGEV LKDIETACKL LNITADPMDW SPSNVQKWLL WTEHQYRLPP MGKAFQELAG   180
KELCAMSEEQ FRQRSPLGGD VLHAHLDIWK SAAWMKERTS PGAIHYCAST SEESWTDSEV   240
DSSCSGQPIH LWQFLKELLL KPHSYGRFIR WLNKEKGIFK IEDSAQVARL WGIRKNRPAM   300
NYDKLSRSIR QYYKKGIIRK PDISQRLVYQ FVHPI
```

SEQ ID NO:101 PEN3 DNA SEQUENCE
Nucleic Acid Accession #:   NM_000742
Coding sequence: 555-2144 (underlined sequences correspond to start and stop codons)

337

```
     1          11         21         31         41         51
     |          |          |          |          |          |
GAGAGAACAG CGTGAGCCTG TGTGCTTGTG TGCTGAGCCC TCATCCCCTC CTGGGGCCAG    60
GCTTGGGTTT CACCTGCAGA ATCGCTTGTG CTGGGCTGCC TGGGCTGTCC TCAGTGGCAC   120
CTGCATGAAG CCGTTCTGGC TGCCAGAGCT GGACAGCCCC AGGAAAACCC ACCTCTCTGC   180
AGAGCTTGCC CAGCTGTCCC CGGGGAAGCCA AATGCCTCTC ATGTAAGTCT TCTGCTCGAC   240
GGGGTGTCTC CTAAACCCTC ACTCTTCAGC CTCTGTTTGA CCATGAAATG AAGTGACTGA   300
GCTCTATTCT GTACCTGCCA CTCTATTTCT GGGGTGACTT TTGTCAGCTG CCCAGAATCT   360
CCAAGCCAGG CTGGTTCTCT GCATCCTTTC AATGACCTGT TTTCTTCTGT AACCACAGGT   420
TCGGTGGTGA GAGGAAGCCT CGCAGAATCC AGCAGAATCC TCACAGAATC CAGCAGCAGC   480
TCTGCTGGGG ACATGGTCCA TGGTGCAACC CACAGCAAAG CCCTGACCTG ACCTCCTGAT   540
GCTCAGGAGA AGCCATGGGC CCCTCCTGTC CTGTGTTCCT GTCCTTCACA AAGCTCAGCC   600
TGTGGTGGCT CCTTCTGACC CCAGCAGGTG GAGAGGAAGC TAAGCGCCCA CCTCCCAGGG   660
CTCCTGGAGA CCCACTCTCC TCTCCCAGTC CCACGGCATT GCCGCAGGGA GGCTCGCATA   720
CCGAGACTGA GGACCGGCTC TTCAAACACC TCTTCCGGGG CTACAACCGC TGGGCGCGCC   780
CGGTGCCCAA CACTTCAGAC GTGGTGATTG TGCGCTTTGG ACTGTCCATC GCTCAGCTCA   840
TCGATGTGGA TGAGAAGAAC CAAATGATGA CCACCAACGT CTGGCTAAAA CAGGAGTGGA   900
GCGACTACAA ACTGCGCTGG AACCCCGCTG ATTTTGGCAA CATCACATCT CTCAGGGTCC   960
CTTCTGAGAT GATCTGGATC CCCGACATTG TTCTCTACAA CAATGCAGAT GGGGAGTTTG  1020
CAGTGACCCA CATGACCAAG GCCCACCTCT TCTCCACGGG CACTGTGCAC TGGGTGCCCC  1080
CGGCCATCTA CAAGAGCTCC TGCAGCATCG ACGTCACCTT CTTCCCCTTC GACCAGCAGA  1140
ACTGCAAGAT GAAGTTTGGC TCCTGGACTT ATGACAAGGC CAAGATCGAC CTGGAGCAGA  1200
TGGAGCAGAC TGTGGACCTG AAGGACTACT GGGAGAGCGG CGAGTGGGCC ATCGTCAATG  1260
CCACGGGCAC CTACAACAGC AAGAAGTACG ACTGCTGCGC CGAGATCTAC CCCGACGTCA  1320
CCTACGCCTT CGTCATCCGG CGGCTGCCCG TCTTCTACAC CATCAACCTC ATCATCCCCT  1380
GCCTGCTCAT CTCCTGCCTC ACTGTGCTGG TCTTCTACCT GCCCTCCGAC TGCGGCGAGA  1440
AGATCACGCT GTGCATTTCG GTGCTGCTGT CACTCACCGT CTTCCTGCTG CTCATCACTG  1500
AGATCATCCC GTCCACCTCG CTGGTCATCC CGCTCATCGG CGAGTACCTG CTGTTCACCA  1560
TGATCTTCGT CACCCTGTCC ATCGTCATCA CCGTCTTCGT GCTCAATGTG CACCACCGCT  1620
CCCCCAGCAC CCACACCATG CCCCACTGGG TGCGGGGGGC CCTTCTGGGC TGTGTGCCCC  1680
GGTGGCTTCT GATGAACCGG CCCCCCACCAC CCGTGGAGCT CTGCCACCCC CTACGCCTGA  1740
AGCTCAGCCC CTCTTATCAC TGGCTGGAGA GCAACGTGGA TGCCGAGGAG AGGGAGGTGG  1800
TGGTGGAGGA GGAGGACAGA TGGGCATGTG CAGGTCATGT GGCCCCCTCT GTGGGCACCC  1860
TCTGCAGCCA CGGCCACCTG CACTCTGGGG CCTCAGGTCC CAAGGCTGAG GCTCTGCTGC  1920
AGGAGGGTGA GCTGCTGCTA TCACCCCACA TGCAGAAGGC ACTGGAAGGT GTGCACTACA  1980
TTGCCGACCA CCTGCCGGTCT GAGGATGCTG ACTCTTCGGT GAAGGAGGAC TGGAAGTATG  2040
TTGCCATGGT CATCGACAGG ATCTTCCTCT GGCTGTTTAT CATCGTCTGC TTCCTGGGGA  2100
CCATCGGCCT CTTTCTGCCT CCGTTCCTAG CTGGAATGAT CTGACTGCAC CTCCCTCGAG  2160
CTGGCTCCCA GGGCAAAGGG GAGGGTTCTT GGATGTGGAA GGGCTTTGAA CAATGTTTAG  2220
ATTTGGAGAT GAGCCCAAAG TGCCAGGGAG AACAGCCAGG TGAGGTGGGA GGTTGGAGAG  2280
CCAGGTGAGG TCTCTCTAAG TCAGGCTGGG GTTGAAGTTT GGAGTCTGTC CGAGTTTGCA  2340
GGGTGCTGAG CTGTATGGTC CAGCAGGGGA GTAATAAGGG CTCTTCCGGA AGGGGAGGAA  2400
GCGGGAGGCA GGCCTGCACC TGATGTGGAG GTACAGGCAG ATCTTCCCTA CCGGGGAGGG  2460
ATGGATGGTT GGATACAGGT GGCTGGGCTA TTCCATCCAT CTGGAAGCAC ATTTGAGCCT  2520
CCAGGCTTCT CCTTGACGTC ATTCCTCTCC TTCCTTGCTG CAAAATGGCT CTGCACCAGC  2580
CGGCCCCCAG GAGGTCTGGC AGAGCTGAGA GCCATGGCCT GCAGGGGCTC CATATGTCCC  2640
TACGCGTGCA GCAGGCAAAC AAGA
```

SEQ ID NO:102 PEN3 Protein sequence
Protein Accession #:   NP_000733

```
     1          11         21         31         41         51
     |          |          |          |          |          |
MGPSCPVFLS FTKLSLWWLL LTPAGGEEAK RPPPRAPGDP LSSPSPTALP QGGSHTETED    60
RLFKHLFRGY NRWARPVPNT SDVVIVRFGL SIAQLIDVDE KNQMMTTNVW LKQEWSDYKL   120
RWNPADFGNI TSLRVPSEMI WIPDIVLYNN ADGEFAVTHM TKAHLFSTGT VHWVPPAIYK   180
SSCSIDVTFF PFDQQNCKMK FGSWTYDKAK IDLEQMEQTV DLKDYWESGE WAIVNATGTY   240
NSKKYDCCAE IYPDVTYAFV IRRLPLFYTI NLIIPCLLIS CLTVLVFYLP SDCGEKITLC   300
ISVLLSLTVF LLLITEIIPS TSLVIPLIGE YLLFTMIFVT LSIVITVFVL NVHHRSPSTH   360
TMPHWVRGAL LGCVPRWLLM NRPPPPVELC HPLRLKLSPS YHWLESNVDA EEREVVVEEE   420
DRWACAGHVA PSVGTLCSHG HLHSGASGPK AEALLQEGEL LLSPHMQKAL EGVHYIADHL   480
RSEDADSSVK EDWKYVAMVI DRIFLWLFII VCFLGTIGLF LPPFLAGMI
```

SEQ ID NO:103 PEU4 DNA SEQUENCE

Nucleic Acid Accession #:   NM_018670
Coding sequence: 87-893 (underlined sequences correspond to start and stop codons)

```
     1          11         21         31         41         51
     |          |          |          |          |          |
CACGAGGCTG GAAGGGGCCA CTTCACACCT CGGGCTCGGC ATAAAGCGGC CGCCGGCCGC    60
CGGCCCCCAG ACGCGCCGCC GCTGCCATGG CCCAGCCCCT GTGCCCGCCG CTCTCCGAGT   120
CCTGGATGCT CTCTGCGGCC TGGGGCCCAA CTCGGCGGCC GCCGCCCTCC GACAAGGACT   180
GCGGCCGCTC CCTCGTCTCG TCCCCAGACT CATGGGGCAG CACCCCAGCC GACAGCCCCG   240
TGGCGAGCCC CGCGCGGCCA GGCACCCTCC GGGACCCCG GGCCCCCTCC GTAGGTAGGC   300
GCGGCGCGCG CAGCAGCCGC CTGGGCAGCG GGCAGAGGCA GAGCGCCAGT GAGCGGGAGA   360
AACTGCGCAT GCGCACGCTG GCCCGCGCCC TGCACGAGCT GCGCCGCTTT CTACCGCCGT   420
CCGTGGCGCC CGCGGGCCAG AGCCTGACCA AGATCGAGAC GCTGCGCCTG GCTATCCGCT   480
ATATCGGCCA CCTGTCGGCC GTGCTAGGCC TCAGCGAGGA GAGTCTCCAG CGCCGGTGCC   540
GGCAGCGCGG TGACGCGGGG TCCCCTCGGG GCTGCCCGCT GTGCCCCGAC GACTGCCCCG   600
CGCAGATGCA GACACGGACG CAGGCTGAGG GGCAGGGGCA GGGGCGCGGG CTGGGCCTGG   660
```

```
        TATCCGCCGT CCGCGCCGGG GCGTCCTGGG GATCCCCGCC TGCCTGCCCC GGAGCCCGAG   720
        CTGCACCCGA GCCGCGCGAC CCGCCTGCGC TGTTCGCCGA GGCGGCGTGC CCGGAAGGGC   780
        AGGCGATGGA GCCAAGCCCA CCGTCCCCGC TCCTTCCGGG CGACGTGCTG GCTCTGTTGG   840
        AGACCTGGAT GCCCCTCTCG CCTCTGGAGT GGCTGCCTGA GGAGCCCAAG TGACAAGGGA   900
        CAACTGACGC CGTCTCTGTG AGCACCGAGG CTTTTTGGCC TCAGCACCTT CGAAGTGGTT   960
        CCTTGGCAGA CTGCCTTTCC TGGAAGAGGG CACGGGCGAT CCCGACGGGG GCATTCCTGC  1020
        GGGTGAGAGC CGTCCCCACC GCGGCGGCCC TTCTCAGCCC CTCCCTCCAT GGAGGGACCC  1080
        ATAGGGCTAG ACACTTTGAG GCAAGCAGGA GGCTCTGCCT AATGTGAATT TATTTATTTG  1140
        TGAATAAACT GTACTGGTGT CAAAAAAAAA AAAAAAAAAA A                      1140
```

SEQ ID NO:104 PEU4 Protein sequence
Protein Accession #:   NP_061140

```
        1          11          21          31          41          51
        |          |           |           |           |           |
        MAQPLCPPLS ESWMLSAAWG PTRRPPPSDK DCGRSLVSSP DSWGSTPADS PVASPARPGT    60
        LRDPRAPSVG RRGARSSRLG SGQRQSASER EKLRMRTLAR ALHELRRFLP PSVAPAGQSL   120
        TKIETLRLAI RYIGHLSAVL GLSEESLQRR CRQRGDAGSP RGCPLCPDDC PAQMQTRTQA   180
        EGQGQGRGLG LVSAVRAGAS WGSPPACPGA RAAPEPRDPP ALFAEAACPE GQAMEPSPPS   240
        PLLPGDVLAL LETWMPLSPL EWLPEEPK
```

SEQ ID NO:105 PEU5 DNA SEQUENCE
Nucleic Acid Accession #:   NM_017636
Coding sequence: 324-3374 (underlined sequences correspond to start and stop codons)

```
        1          11          21          31          41          51
        |          |           |           |           |           |
        CCACGGAGAA GCCCACCGAT GCCTACGGAG AGCTGGACTT CACGGGGGCC GGCCGCAAGC    60
        ACAGCAATTT CCTCCGGCTC TCTGACCGAA CGGATCCAGC TGCAGTTTAT AGTCTGGTCA   120
        CACGCACATG GGGCTTCCGT GCCCCGAACC TGGTGGTGTC AGTGCTGGGG GGATCGGGGG   180
        GCCCCGTCCT CCAGACCTGG CTGCAGGACC TGCTGCGTCG TGGGCTGGTG CGGGCTGCCC   240
        AGAGCACAGG AGCCTGGATT GTCACTGGGG GTCTGCCACG GGGCATCGGC CGGCATGTTG   300
        GTGTGGCTGT ACGGGACCAT CAGATGGCCA GCACTGGGGG CACCAAGGTG GTGGCCATGG   360
        GTGTGGCCCC CTGGGGTGTG GTCCGGAATA GAGACACCCT CATCAACCCC AAGGGCTCGT   420
        TCCCTGCGAG GTACCGGTGG CGCGGTGACC CGGAGGACGG GGTCCAGTTT CCCCTGGACT   480
        ACAACTACTC GGCCTTCTTC CTGGTGGACG ACGGCACACA CGGCTGCCTG GGGGGCGAGA   540
        ACCGCTTCCG CTTGCGCCTG GAGTCCTACA TCTCACAGCA GAAGACGGGC GTGGGAGGGA   600
        CTGGAATTGA CATCCCTGTC CTGCTCCTCC TGATTGATGG TGATGAGAAG ATGTTGACGC   660
        GAATAGAGAA CGCCACCCAG GCTCAGCTCC CATGTCTCCT CGTGGCTGGC TCAGGGGGAG   720
        CTGCGGACTG CCTGGCGGAG ACCCTGGAAG ACACTCTGGC CCCAGGGAGT GGGGGAGCCA   780
        GGCAGGCCGA AGCCCGAGAT CGAATCAGGC GTTTCTTTCC CAAAGGGGAC CTTGAGGTCC   840
        TGCAGGCCCA GGTGGAGAGG ATTATGACCC GGAAGGAGCT CCTGACAGTC TATTCTTCTG   900
        AGGATGGGTC TGAGGAATTC GAGACCATAG TTTTGAAGGC CCTTGTGAAG GCCTGTGGGA   960
        GCTCGGAGGC CTCAGCCTAC CTGGATGAGC TGCGTTTGGC TGTGGCTTGG AACCGCGTGG  1020
        ACATTGCCCA GAGTGAACTC TTTCGGGGGG ACATCCAATG GCGGTCCTTC CATCTCGAAG  1080
        CTTCCCTCAT GGACGCCCTG CTGAATGACC GGCCTGAGTT CGTGCGCTTG CTCATTTCCC  1140
        ACGGCCTCAG CCTGGGCCAC TTCCTGACCC CGATGCGCCT GGCCCAACTC TACAGCGCGG  1200
        CGCCCTCCAA CTCGCTCATC CGCAACCTTT TGGACCAGGC GTCCCACAGC GCAGGCACCA  1260
        AAGCCCCAGC CCTAAAAGGG GGAGCTGCGG AGCTCCGGCC CCCTGACGTG GGGCATGTGC  1320
        TGAGGATGCT GCTGGGGAAG ATGTGCGCGC CGAGGTACCC CTCCGGGGGC GCCTGGGACC  1380
        CTCACCCAGG CCAGGGCTTC GGGGAGAGCA TGTATCTGCT CTCGGACAAG GCCACCTCGC  1440
        CGCTCTCGCT GGATGCTGGC CTCGGGCAGG CCCCCTGGAG CGACCTGCTT CTTTGGGCAC  1500
        TGTTGCTGAA CAGGGCACAG ATGGCCATGT ACTTCTGGGA GATGGGTTCC AATGCAGTTT  1560
        CCTCAGCTCT TGGGGCCTGT TTGCTGCTCC GGGTGATGGC ACGCGTGGAG CCTGACGCTG  1620
        AGGAGGCAGC ACGGAGGAAA GACCTGGCGT TCAAGTTTGA GGGGGATGGG C GTTGACCTCT  1680
        TTGGCGAGTG CTATCGCAGC AGTGAGGTGA GGGCTGCCCG CCTCCTCCTC CGTCGCTGCC  1740
        CGCTCTGGGG GGATGCCACT TGCCTCCAGC TGGCCATGCA AGCTGACGCC CGTGCCTTCT  1800
        TTGCCCAGGA TGGGGTACAG TCTCTGCTGA CACAGAAGTG GTGGGGAGAT ATGGCCAGCA  1860
        CTACACCCAT CTGGGCCCTG GTTCTCGCCT TCTTTTGCCC TCCACTCATC TACACCCGCC  1920
        TCATCACCTT CAGGAAATCA GAAGAGGAGC CCACACGGGA GGAGCTAGAG TTTGACATGG  1980
        ATAGTGTCAT TAATGGGGAA GGGCCTGTCG GGACGGCGGA CCCAGCCGAG AAGACGCCGC  2040
        TGGGGGTCCC GCGCCAGTCG GGCCGTCCGG GTTGCTGCGG GGGCCGCTGC GGGGGGCGCC  2100
        GGTGCCTACG CCGCTGGTTC CACTTCTGGG GCGCGCCGGT GACCATCTTC ATGGGCAACG  2160
        TGGTCAGCTA CCTGCTGTTC TTGCTGCTTT TCTCGCGGGT GCTGCTCGTG GATTTCCAGC  2220
        CGGCGCCGCC CGGCTCCCTG GAGCTGCTGC TCTATTTCTG GGCTTTCACG CTGCTGTGCG  2280
        AGGAACTGCG CCAGGGCCTG AGCGGGAGGCG GGGGCAGCCT CGCCAGCGGG GGCCCCGGGC  2340
        CTGGCCATGC CTCACTGAGC CAGCGCCTGC GCCTCTACCT CGCCGACAGC TGGAACCAGT  2400
        GCGACCTAGT GGCTCTCACC TGCTTCCTCC TGGGCGTGGG CTGCCGGCTG ACCCCGGGTT  2460
        TGTACCACCT GGGCCGCACT GTCCTCTGCA TCGACTTCAT GGTTTTCACG GTGCGGCTGC  2520
        TTCACATCTT CACGGTCAAC AAACAGCTGG GGCCCAAGAT CGTCATCGTG AGCAAGATGA  2580
        TGAAGGACGT GTTCTTCTTC CTCTTCTTCC TCGGCGTGTG GCTGGTAGCC TATGGCGTGG  2640
        CCACGGAGGG GCTCCTGAGG CCACGGGACA GTGACTTCCC AAGTATCCTG CGCCGCGTCT  2700
        TCTACCGTCC CTACCTGCAG ATCTTCGGGC AGATTCCCCA GGAGGACATG GACGTGGCCC  2760
        TCATGGAGCA CAGCAACTGC TCGTCGGAGC CCGGCTTCTG GGCACACCCT CCTGGGGCCC  2820
        AGGCGGGCAC CTGCGCGTCC CAGTATGCCA ACTGGCTGGT GGTGCTGCTC CTCGTCATCT  2880
        TCCTGCTCGT GGCCAACATC CTGCTGGTCA ACTTGCTCAT TGCCATGTTC AGTTACACAT  2940
        TCGGCAAAGT ACAGGGCAAC AGCGATCTCT ACTGGAAGGC GCAGCGTTAC CGCCTCATCC  3000
        GGGAATTCCA CTCTCGGCCC GCGCTGGCCC CGCCCTTTAT CGTCATCTCC CACTTGCGCC  3060
        TCCTGCTCAG GCAATTGTGC AGGCGCACCC GGAGCCCCCA GCCGTCCTCC CCGGCCCTCG  3120
        AGCATTTCCG GGTTTACCTT TCTAAGGAAG CCGAGCGGAA GCTGCTAACG TGGGAATCGG  3180
```

```
TGCATAAGGA GAACTTTCTG CTGGCACGCG CTAGGGACAA GCGGGAGAGC GACTCCGAGC   3240
GTCTGGAGCG CACGTCCCAG AAGGTGGACT TGGCACTGAA ACAGCTGGGA CACATCCGCG   3300
AGTACGAACA GCGCCTGAAA GTGCTGGAGC GGGAGGTCCA GCAGTGTAGC CGCGTCCTGG   3360
GGTGGGTGAC GTAGGCCGTT AGCAGCTCTG CCATGTTGCC CTCAGGTGGG CCGCCACCCC   3420
TTGACCTGCA TGGGTCCAAA GAGTGAGCCA TGCTGGCGGA TTTTAAGGAG AAGCCCCCAC   3480
AGGGGATTTT GCTCTTAGAG TAAGGCTCAT GTGGGCCTCG GCCCCCGCAC CTGGTGGCCT   3540
TGTCCTTGAG GTGAGCCCCA TGTCCATCTG GGCCACTGTC AGGACCACCT TTGGGAGTGT   3600
CATCCTTACA AACCACAGCA TGCCCGGCTC CTCCCAGAAC CAGTCCCAGC CTGGGAGGAT   3660
CAAGGCCTGG ATCCCGGGGC GTTATCCATC TGGAGGCTGC AGGGTCCTTG GGGTAACAGG   3720
GACCACAGAC CCCTCACCAC TCACAGATTC CTCACACTGG GGAAATAAAG CCATTTCAGA   3780
GGAAAAAAAA AAAAAAAAAA AAAAAAAAAA
```

SEQ ID NO:106 PEU5 Protein sequence
Protein Accession #:   NP_060106

```
 1          11         21         31         41         51
 |          |          |          |          |          |
MASTGGTKVV AMGVAPWGVV RNRDTLINPK GSFPARYRWR GDPEDGVQFP LDYNYSAFFL    60
VDDGTHGCLG GENRFRLRLE SYISQQKTGV GGTGIDIPVL LLLIDGDEKM LTRIENATQA   120
QLPCLLVAGS GGAADCLAET LEDTLAPGSG GARQGEARDR IRRFFPKGDL EVLQAQVERI   180
MTRKELLTVY SSEDGSEEFE TIVLKALVKA CGSSEASAYL DELRLAVAWN RVDIAQSELF   240
RGDIQWRSFH LEASLMDALL NDRPEFVRLL ISHGLSLGHF LTPMRLAQLY SAAPSNSLIR   300
NLLDQASHSA GTKAPALKGG AAELRPPDVG HVLRMLLGKM CAPRYPSGGA WDPHPGQGFG   360
ESMYLLSDKA TSPLSLDAGL GQAPWSDLLL WALLLNRAQM AMYFWEMGSN AVSSALGACL   420
LLRVMARLEP DAEEAARRKD LAFKFEGMGV DLFGECYRSS EVRAARLLLR RCPLWGDATC   480
LQLAMQADAR AFFAQDGVQS LLTQKWWGDM ASTTPIWALV LAFFCPPLIY TRLITFRKSE   540
EEPTREELEF DMDSVINGEG PVGTADPAEK TPLGVPRQSG RPGCCGGRCG GRRCLRRWFH   600
FWGAPVTIFM GNVVSYLLFL LLFSRVLLVD FQPAPPGSLE LLLYFWAFTL LCEELRQGLS   660
GGGGSLASGG PGPGHASLSQ RLRLYLADSW NQCDLVALTC FLLGVGCRLT PGLYHLGRTV   720
LCIDFMVFTV RLLHIFTVNK QLGPKIVIVS KMMKDVFFFL FFLGVWLVAY GVATEGLLRP   780
RDSDFPSILR RVFYRPYLQI FGQIPQEDMD VALMEHSNCS SEPGFWAHPP GAQAGTCVSQ   840
YANWLVVLLL VIFLLVANIL LVNLLIAMFS YTFGKVQGNS DLYWKAQRYR LIREFHSRPA   900
LAPPFIVISH LRLLLRQLCR RPRSPQPSSP ALEHFRVYLS KEAERKLLTW ESVHKENFLL   960
ARARDKRESD SERLERTSQK VDLALKQLGH IREYEQRLKV LEREVQQCSR VLGWVT
```

                                    SEQ ID NO:107 PEW3 DNA SEQUENCE

Nucleic Acid Accession #:   NM_005982
Coding sequence: 276-1130 (underlined sequences correspond to start and stop codons)

```
 1          11         21         31         41         51
 |          |          |          |          |          |
GGTAGCAGCA TCCACCGGGC GGGAGGTCGG AGGCAGCAAG GCCTTAAAGG CTACTGAGTG    60
CGCCGGCCGT TCCGTGTCCA GAACCTCCCC TACTCCTCCG CCTTCTCTTC CTTGGCCGCC   120
CACCGCCAAG TTCCGACTCC GGTTTTCGCC TTTGCAAAGC CTAAGGAGGA GGTTAGGAAC   180
AGCCGCGCCC CCCTCCCTGC GGCCGCCGCC CCCTGCCTCT CGGCTCTGCT CCCTGCCGCG   240
TGCGCCTAGG CCGTGCGCCC CGGCAGGCGC CAGCCATGTC GATGCTGCCG TCGTTTGGCT   300
TTACGCAGGA GCAAGTGGCG TGCGTGTGCG AGGTTCTGCA GCAAGGCGGA AACCTGGAGC   360
GCCTGGGCAG GTTCCTGTGG TCACTGCCCG CCTGCGACCA CCTGCACAAG AACGAGAGCG   420
TACTCAAGGC CAAGGCGGTG GTCGCCTTCC ACCGCGGCAA CTTCCGTGAG CTCTACAAGA   480
TCCTGGAGAG CCACCAGTTC TCGCCTCACA ACCACCCCAA ACTGCAGCAA CTGTGGCTGA   540
AGGCGCATTA CGTGGAGGCC GAGAAGCTGC GCGGCCGACC CCTGGGCGCC GTGGGCAAAT   600
ATCGGGTGCG CCGAAAATTT CCACTGCCGC GCACCATCTG GGACGGCGAG GAGACCAGCT   660
ACTGCTTCAA GGAGAAGTCG AGGGGTGTCC TGCGGGAGTG GTACGCGCAC AATCCCTACC   720
CATCGCCGCG TGAGAAGCGG GAGCTGGCCG AGGCCACCGG CCTCACCACC ACCCAGGTCA   780
GCAACTGGTT TAAGAACCGG AGGCAAAGAG ACCGGGCCGC GGAGGCCAAG GAAAGGGAAA   840
ACACCGAAAA CAATAACTCC TCCTCCAACA AGCAGAACCA ACTCTCTCCT CTGGAAGGGG   900
GCAAGCCGCT CATGTCCAGC TCAGAAGAGG AATTCTCACC TCCCCAAAGT CCAGACCAGA   960
ACTCGGTCCT TCTGCTGCAG GGCAATATGG GCCACGCCAG GAGCTCAAAC TATTCTCTCC  1020
CGGGCTTAAC AGCCTCGCAG CCCAGTCACG GCCTGCAGAC CCACCAGCAT CAGCTCCAAG  1080
ACTCTCTGCT CGGCCCCCTC ACCTCCAGTC TGGTGGACTT GGGGTCCTAA GTGGGGAGGG  1140
ACTGGGGCCT CGAAGGGATT CCTGGAGCAG CAACCACTGC AGCGACTAGG GACACTTGTA  1200
AATAGAAATC AGGAACATTT TTGCAGCTTG TTTCTGGAGT TGTTTGCGCA TAAAGGAATG  1260
GTGGACTTTC ACAAATATCT TTTTAAAAAT CAAAACCAAC AGCGATCTCA AGCTTAATCT  1320
CCTCTTCTCT CCAACTCTTT CCACTTTTGC ATTTTCCTTC CCAATGCAGA GATCAGGG
```

SEQ ID NO:108 PEW3 Protein sequence
Protein Accession #:   NP_005973

```
 1          11         21         31         41         51
 |          |          |          |          |          |
MSMLPSFGFT QEQVACVCEV LQQGGNLERL GRFLWSLPAC DHLHKNESVL KAKAVVAFHR    60
GNFRELYKIL ESHQFSPHNH PKLQQLWLKA HYVEAEKLRG RPLGAVGKYR VRRKFPLPRT   120
IWDGEETSYC FKEKSRGVLR EWYAHNPYPS PREKRELAEA TGLTTTQVSN WFKNRRQRDR   180
AAEAKERENT ENNNSSSNKQ NQLSPLEGGK PLMSSSEEEF SPPQSPDQNS VLLLQGNMGH   240
ARSSNYSLPG LTASQPSHGL QTHQHQLQDS LLGPLTSSLV DLGS
```

                                    SEQ ID NO:109 PFJ8 DNA SEQUENCE

Nucleic Acid Accession #:   NM_005069
Coding sequence: 57-2060 (underlined sequences correspond to start and stop codons)

340

```
        1     11    21    31    41    51
        |     |     |     |     |     |
   GGGGCTCCGC GGGCCTGGAG CACGGCCGGG TCTAATATGC CCGGAGCCGA GGCGCGATGA    60
   AGGAGAAGTC CAAGAATGCG GCCAAGACCA GGAGGGAGAA GGAAAATGGC GAGTTTTACG   120
   AGCTTGCCAA GCTGCTCCCG CTGCCGTCGG CCATCACTTC GCAGCTGGAC AAAGCGTCCA   180
   TCATCCGCCT CACCACGAGC TACCTGAAGA TGCGCGCCGT CTTCCCCGAA GGTTTAGGAG   240
   ACGCGTGGGG ACAGCCGAGC CGCGCCGGGC CCCTGGACGG CGTCGCCAAG GAGCTGGGAT   300
   CGCACTTGCT GCAGACTTTG GATGGATTTG TTTTTGTGGT AGCATCTGAT GGCAAAATCA   360
   TGTATATATC CGAGACCGCT TCTGTCCATT TAGGCTTATC CCAGGTGGAG CTCACGGGCA   420
   ACAGTATTTA TGAATACATC CATCCTTCTG ACCACGATGA GATGACCGCT GTCCTCACGG   480
   CCCACCAGCC GCTGCACCAC CACCTGCTCC AAGAGTATGA GATAGAGAGG TCGTTCTTTC   540
   TTCGAATGAA ATGTGTCTTG GCGAAAAGGA ACGCGGGCCT GACCTGCAGC GGATACAAGG   600
   TCATCCACTG CAGTGGCTAC TTGAAGATCA GGCAGTATAT GCTGGACATG TCCCTGTACG   660
   ACTCCTGCTA CCAGATTGTG GGGCTGGTGG CCGTGGGCCA GTCGCTGCCA CCCAGTGCCA   720
   TCACCGAGAT CAAGCTGTAC AGTAACATGT TCATGTTCAG GGCCAGCCTT GACCTGAAGC   780
   TGATATTCCT GGATTCCAGG GTGACCGAGG TGACGGGTTA CGAGCCGCAG GACCTGATCG   840
   AGAAGACCCT ATACCATCAC GTGCACGGCT GCGACGTGTT CCACCTCCGC TACGCACACC   900
   ACCTCCTGTT GGTGAAGGGC CAGGTCACCA CCAAGTACTA CCGGCTGCTG TCCAAGCGGG   960
   GCGGCTGGGT GTGGGTGCAG AGCTACGCCA CCGTGGTGCA CAACAGCCGC TCGTCCCGGC  1020
   CCCACTGCAT CGTGAGTGTC AATTATGTAC TCACGGGAGT TGAATACAAG GAACTTCAGC  1080
   TGTCCCTGGA GCAGGTGTCC ACTGCCAAGT CCCAGGACTC CTGGAGGACC GCCTTGTCTA  1140
   CCTCACAAGA AACTAGGAAA TTAGTGAAAC CCAAAAATAC CAAGATGAAG ACAAAGCTGA  1200
   GAACAAACCC TTACCCCCCA CAGCAATACA GCTCGTTCCA AATGGACAAA CTGGAATGCG  1260
   GCCAGCTCGG AAACTGGAGA GCCAGTCCCC CTGCAAGCGC TGCTGCTCCT CCAGAACTGC  1320
   AGCCCCACTC AGAAAGCAGT GACCTTCTGT ACACGCCATC CTACAGCCTG CCCTTCTCCT  1380
   ACCATTACGG ACACTTCCCT CTGGACTCTC ACGTCTTCAG CAGCAAAAAG CCAATGTTGC  1440
   CGGCCAAGTT CGGGCAGCCC CAAGGATCCC CTTGTGAGGT GGCACGCTTT TTCCTGAGCA  1500
   CACTGCCAGC CAGCGGTGAA TGCCAGTGGC ATTATGCCAA CCCCCTAGTG CCTAGCAGCT  1560
   CGTCTCCAGC TAAAAATCCT CCAGAGCCAC CGGCGAACAC TGCTAGGCAC AGCCTGGTGC  1620
   CAAGCTACGA AGCGCCCGCC GCCGCCGTGC GCAGGTTCGG CGAGGACACC GCGCCCCCGA  1680
   GCTTCCCGAG CTGCGGCCAC TACCGCGGAG AGCCCGGGCC TGGGCCGGCC AAAGCCGCCC  1740
   GCCAGGCCGC CCGGGACGGG GCGCGGCTGG CGCTGGCCCG CGCGGCACCC GAGTGCTGCG  1800
   CGCCCCCGAC CCCCGAGGCC CCGGGCGCGC CGGCGCAGCT GCCCTTCGTG CTGCTCAACT  1860
   ACCACCGCGT GCTGGCCCGG CGCGGACCGC TGGGGGGCGC CGCACCCGCC GCCTCCGGCC  1920
   TGGCCTGCGC TCCCGGCGGC CCCGAGGCGG CGACCGGCGC GCTGCGGCTC CGGCACCCGA  1980
   GCCCCGCCGC CACCTCCCCG CCCGGCCGCG CCCTGCCGCA CTACCTGGGC GCCTCGGTCA  2040
   TCATCACCAA CGGGAGGTGA CCCGCTGGCC GCCCGCGCCA GGAGCCTGGA CCCGGCCTCC  2100
   CGGGGCTGCG GCGCGCCACCGA GCCCGGCAAA TGCGCACGAC CTACATTAAT TTATGCAGAG  2160
   ACAGCTGTTT GAATTGGACC CCGCCGCCGA CTTGCGGATT CCACCGCGG AGGCCCCGCG   2220
   CGCCGGTGCC GAGGGCCGAG GAGCGCCCGG GTCCGGGCAG GTGACCGCCC GCCTCTGTCC  2280
   TGCGAGGGCC GGTGCGACCC AGTTGCTGGG GGCTTGGTTT CCTCACCTTG AAATCGGGCT  2340
   TCACGCGTCT TGCCTTGTCC CCAACGTTCC ACAACAGTCC CGCTGGGGGA TTGAAGCGGT  2400
   TTCACTCCGC AAATATCCTC CACTTTCAGG AGGGAAAACC CACCCTACCA CAGTCCGCTC  2460
   TTCCAAGTGG ACGGCAGACC TGGGAGGGGA CGCCTGTGTC ACGAGCCCTT TTAGATGCTT  2520
   AGGTGAAGGC AGAAGTGATG ATTGTAAGTC CCATGAATAC ACAACTCCAC TGTCTTTAAA  2580
   AGTCATTCAA GAGTCTCATT ATTTTTGTTT TTATTTAACC CTTTCTTCAA TACAAAAAGC  2640
   CAACAAACCA AGACTAAGGG GGTGACCATG CAATTCCATT TTGTGTCTGT GAACATAGGT  2700
   GTGCTTCCCA AATACATTAA CAAGCTCTTA CTTCCCCCTA ACCCCTATGA ACTCTTGATA  2760
   ACACCAAGAG TAGCACCTTC AGAATATATT GAATAGGCAT TAAATGCAAA AATATATATG  2820
   TAGCCAGACA GTTTATGAGA ATGACCCTGT CAAGCTTCAT TATTACGTGG CAAAATCCCT  2880
   CTGGCCCACA CAGATCTGTA ATTCACTAGG CTCGTGTTTG CTACAAATAG TGCTAATAAA  2940
   GTTAAATTGC ACGTGCAATA CGGAACACTG TCAATGGACT GCACCTTGTG AAGGAAAAAC  3000
   ATGCTTAAGG GGGTGTAATG AAAATGATGT AGACATTTTA AGCATTTTCT ACACAGCGAG  3060
   AAAACTTCGT AAGAACATGT TACGTGTGCA ACAGGTAAAC AGAAATCCTT TCATAAAGCA  3120
   CCAGCAGTGT TTAAAAAATG AGCTTCCATT AATTTTTACT TTTTATGGGT TTTGCTTAAA  3180
   GATCTCAACA TGGAAAAATC CTGTCATGGC TCTGAACTGC ACAATGCATT GAACCGCCGT  3240
   CCTTCAATTT TCTTCACACT ATCAACACTG CAGCATTTTG CTGCTTATC AAAATGGTTT  3300
   ATTTTAGGAA ACTTTTTCCA CCTTTCTGAA TGGAAAGAGG TTTTCACAAA TGTTTTAAAC  3360
   TCATCGTTCT AAAATCAAGT GCACCTACAC CAACTGCCTT CAAAATGTGA ACTGACTTTT  3420
   TTTTTTTTTT TTTTGCCAAC CCTGTGTCAC TTAGTGAGGA CCTGACACAA TCCCTACAGG  3480
   GTGTCTGTCA GTGGGCCTCA TGGTAAGAGT CACAATTTGC AAATTTAGGA CCGTGGGTCA  3540
   TGCAGCGAAG GGGCTGGATG GTAGGAAGGG ATGTGCCCGC CTCTCCACGC ACTCAGCTAT  3600
   ACCTCATTCA CAGCTCCTTG TGAGTGTGTG CACAGGAAAT AAGCCGAGGG TATTATTTTT  3660
   TTATGTTCAT GAGTCTTGTA ATTAAACCGT GATTCTTGAA AGGTGTAGGT TTGATTACTA  3720
   GGAGATACCA CCGACATTTT TCAATAAAGT ACTGCAAAAT GCTTTTGTGT CTACCTTGTT  3780
   ATTAACTTTT GGGGCTGTAT TTAGTAAAAA TAAATCAAGG CTATCGGAGC AGTTCAATAA  3840
   CAAAGGTTAC TGTTGAGAAA AAAGACCCTA TCATAGATTT ACAAG
```

SEQ ID NO:110 PFJ8 Protein sequence:
Protein Accession #:   NP_005060.1

```
        1     11    21    31    41    51
        |     |     |     |     |     |
   MKEKSKNAAK TRREKENGEF YELAKLLPLP SAITSQLDKA SIIRLTTSYL KMRAVFPEGL    60
   GDAWGQPSRA GPLDGVAKEL GSHLLQTLDG FVFVVASDGK IMYISETASV HLGLSQVELT   120
   GNSIYEYIHP SDHDEMTAVL TAHQPLHHHL LQEYEIERSF FLRMKCVLAK RNAGLTCSGY   180
   KVIHCSGYLK IRQYMLDMSL YDSCYQIVGL VAVGQSLPPS AITEIKLYSN MFMFRASLDL   240
```

341

```
KLIFLDSRVT EVTGYEPQDL IEKTLYHHVH GCDVFHLRYA HHLLLVKGQV TTKYYRLLSK   300
RGGWVWVQSY ATVVHNSRSS RPHCIVSVNY VLTEIEYKEL QLSLEQVSTA KSQDSWRTAL   360
STSQETRKLV KPKNTKMKTK LRTNPYPPQQ YSSFQMDKLE CGQLGNWRAS PPASAAAPPE   420
LQPHSESSDL LYTPSYSLPF SYHYGHFPLD SHVFSSKKPM LPAKFGQPQG SPCEVARFFL   480
STLPASGECQ WHYANPLVPS SSSPAKNPPE PPANTARHSL VPSYEAPAAA VRRFGEDTAP   540
PSFPSCGHYR EEPALGPAKA ARQAARDGAR LALARAAPEC CAPPTPEAPG APAQLPFVLL   600
NYHRVLARRG PLGGAAPAAS GLACAPGGPE AATGALRLRH PSPAATSPPG APLPHYLGAS   660
VIITNGR
```

SEQ ID NO:111 PFJ7 DNA SEQUENCE

Nucleic Acid Accession #:   NM_006549
Coding sequence: 1-1254 (underlined sequences correspond to start and stop codons)

```
 1    11   21   31   41   51
 |    |    |    |    |    |
ATGAACGGAC GCTGCATCTG CCCGTCCCTG CCCTACTCAC CCGTCAGCTC CCCGCAGTCC   60
TCGCCTCGGC TGCCCCGGCG GCCGACAGTG GAGTCTCACC ACGTCTCCAT CACGGGTATG   120
CAGGACTGTG TGCAGCTGAA TCAGTATACC CTGAAGGATG AAATTGGAAA GGGCTCCTAT   180
GGTGTCGTCA AGTTGGCCTA CAATGAAAAT GACAATACCT ACTATGCAAT GAAGGTGCTG   240
TCCAAAAAGA AGCTGATCCG GCAGGCCGGC TTTCCACGTC GCCCTCCACC CCGAGGCACC   300
CGGCCAGCTC CTGGAGGCTG CATCCAGCCC AGGGGCCCCA TTGAGCAGGT GTACCAGGAA   360
ATTGCCATCC TCAAGAAGCT GGACCACCCC AATGTGGTGA AGCTGGTGGA GGTCCTGGAT   420
GACCCCAATG AGGACCATCT GTACATGGTG TTCGAACTGG TCAACCAAGG GCCCGTGATG   480
GAAGTGCCCA CCCTCAAACC ACTCTCTGAA GACCAGGCCC GTTTCTACTT CCAGGATCTG   540
ATCAAAGGCA TCGAGTACTT ACACTACCAG AAGATCATCC ACCGTGACAT CAAACCTTCC   600
AACCTCCTGG TCGGAGAAGA TGGGCACATC AAGATCGCTG ACTTTGGTGT GAGCAATGAA   660
TTCAAGGGCA GTGACGCGCT CCTCTCCAAC ACCGTGGGCA CGCCCGCCTT CATGGCACCC   720
GAGTCGCTCT CTGAGACCCG CAAGATCTTC TCTGGGAAGG CCTTGGATGT TTGGGCCATG   780
GGTGTGACAC TATACTGCTT TGTCTTTGGC CAGTGCCCAT TCATGGACGA GCGGATCATG   840
TGTTTACACA GTAAGATCAA GAGTCAGGCC CTGGAATTTC CAGACCAGCC CGACATAGCT   900
GAGGACTTGA AGGACCTGAT CACCCGTATG CTGGACAAGA ACCCCGAGTC GAGGATCGTG   960
GTGCCGGAAA TCAAGCTGCA CCCCTGGGTC ACGAGGCATG GGGCGGAGCC GTTGCCGTCG   1020
GAGGATGAGA ACTGCACGCT GGTCGAAGTG ACTGAAGAGG AGGTCGAGAA CTCAGTCAAA   1080
CACATTCCCA GCTTGGCAAC CGTGATCCTG GTGAAGACCA TGATACGTAA ACGCTCCTTT   1140
GGGAACCCAT TCGAGGGCAG CCGGCGGGAG GAACGCTCAC TGTCAGCGCC TGGAAACTTG   1200
CTCACCAAAA AACCAACCAG GGAAATGTGAG TCCCTGTCTG AGCTCAAGAC CTAGAAAATA   1260
AGTCCCCTTC CTGCCTGTTG CAAAGTAACG TAAGAGTTCC CTCACCCGAG TGGATGCAGA   1320
CGTTCTTGCT GTCAGCCACC TTCCTTCATA CACATAGCCA GCCCAGGGTG ACCAGAACGT   1380
CCCAGGACAG ATGAGGCTTT GTGTCCTTAT GAGAGTGGGA GAACCTGGTG GGCACCCCTG   1440
GTGCAGGTGC TGTGGTGGGT GGGGACCCCA CTGCCTTTCC CACTGAGCAC ATCATGGCTA   1500
CCTGACTTGG TGGGAGTTCC ATTCAGTCAC TTCTGTTTCT TAAACATAGC TTTACTGAGG   1560
TACAATTCAC ATACCATGTA ATTCACCCAC GGGAAGTGTA TGATTCAGTG GTTTCTAATA   1620
CACACTTCTG CAGCCATTAC CACCGTCAAC TTTACGACAT TTTCATCAGC CCAAGAAGAC   1680
ACCCTACACT CCTTAGCTGT CCCCATCCAA CTCCCCCACC CCAGTAACCA CTCAGAATAG   1740
GTATGGATTT GCCTATTCTG GACGTTTCGT ATAAATGGCG TCATACACTA AAAAAAAAAA   1800
AAAA
```

SEQ ID NO:112 PFJ7 Protein sequence:
Protein Accession #:   NP_006540.1

```
 1    11   21   31   41   51
 |    |    |    |    |    |
MNGRCICPSL PYSPVSSPQS SPRLPRRPTV ESHHVSITGM QDCVQLNQYT LKDEIGKGSY   60
GVVKLAYNEN DNTYYAMKVL SKKKLIRQAG FPRRPPPRGT RPAPGGCIQP RGPIEQVYQE   120
IAILKKLDHP NVVKLVEVLD DPNEDHLYMV FELVNQGPVM EVPTLKPLSE DQARFYFQDL   180
IKGIEYLHYQ KIIHRDIKPS NLLVGEDGHI KIADFGVSNE FKGSDALLSN TVGTPAFMAP   240
ESLSETRKIF SGKALDVWAM GVTLYCFVFG QCPFMDERIM CLHSKIKSQA LEFPDQPDIA   300
EDLKDLITRM LDKNPESRIV VPEIKLHPWV TRHGAEPLPS EDENCTLVEV TEEEVENSVK   360
HIPSLATVIL VKTMIRKRSF GNPFEGSRRE ERSLSAPGNL LTKKPTRECE SLSELKT
```

SEQ ID NO:113 PFJ6 DNA SEQUENCE

Nucleic Acid Accession #:   NM_021810
Coding sequence: 1-429 (underlined sequences correspond to start and stop codons)

```
 1    11   21   31   41   51
 |    |    |    |    |    |
ATGAAACCTC TGATATGGAC ATGGTCAGAT GTTGAAGGCC AGAGGCCGGC TCTGCTCATC   60
TGCACAGCTG CAGCAGGACC CACGCAGGGA GTTAAGGGTT ATGGCAAGCC CTTTGAGCCA   120
AGAAGTGTGA AAAACATACA CTCTACTCCT GCTTACCCAG ATGCCACAAT GCACAGACAA   180
CTCCTGGCTC CGGTGGAAGG AAGGATGGCA GAGACATTGA ATCAGAAACT CCATGTTGCC   240
AATGTGCTGG AAGATGACCC CGGCTACCTA CCTCACGTCT ACAGCGAGGA AGGGGGAGTGT   300
GGAGGGGCCC CATCCCTCAG CTCTCTGGCC AGCTTGGAAC AGGAGTTGCA ACCTGATTTG   360
```

CTGGACTCTT TGGGTTCAAA AGCGACTCCG TTTGAGGAAA TATATTCAGA GTCAGGTGTT 420
CCTTCC<u>TAA</u>

5

SEQ ID NO:114 <u>PFJ6 Protein sequence:</u>
Protein Accession #:   NP_068582.1

```
1    11    21    31    41    51
|    |     |     |     |     |
|    |     |     |     |     |
```
10
MKPLIWTWSD VEGQRPALLI CTAAAGPTQG VKGYGKPFEP RSVKNIHSTP AYPDATMHRQ 60
LLAPVEGRMA ETLNQKLHVA NVLEDDPGYL PHVYSEEGEC GGAPSLSSLA SLEQELQPDL 120
LDSLGSKATP FEEIYSESGV PS

15

<div align="center">SEQ ID NO:115 PFJ5 DNA SEQUENCE</div>

Nucleic Acid Accession #:   NM_006361
Coding sequence: 131-985 (underlined sequences correspond to start and stop codons)

20
```
1    11    21    31    41    51
|    |     |     |     |     |
|    |     |     |     |     |
```
CGAATGCAGG CGACTTGCGA GCTGGGAGCG ATTTAAAACG CTTTGGATTC CCCCGGCCTG 60
GGTGGGGAGA GCGAGCTGGG TGCCCCCTAG ATTCCCCGCC CCCGCACCTC ATGAGCCGAC 120
CCTCGGCTCC <u>ATG</u>GAGCCCG GCAATTATGC CACCTTGGAT GGAGCCAAGG ATATCGAAGG 180
CTTGCTGGGA GCGGGAGGGG GGCGGAATCT GGTCGCCCAC TCCCCTCTGA CCAGCCACCC 240
AGCGGCGCCT ACGCTGATGC CTGCTGTCAA CTATGCCCCC TTGGATCTGC CAGGCTCGGC 300
GGAGCCGCCA AAGCAATGCC ACCCATGCCC TGGGGTGCCC CAGGGGACGT CCCCAGCTCC 360
CGTGCCTTAT GGTTACTTTG GAGGCGGGTA CTACTCCTGC CGAGTGTCCC GGAGCTCGCT 420
GAAACCCTGT GCCCAGGCAG CCACCCTGGC CGCGTACCCC GCGGAGACTC CCACGGCCGG 480
GGAAGAGTAC CCCAGTCGCC CCACTGAGTT TGCCTTCTAT CCGGGGATATC CGGGAACCTA 540
CCACGCTATG GCCAGTTACC TGGACGTGTC TGTGGTGCAA ACTCTGGGTG CTCCTGGAGA 600
ACCGCGACAT GACTCCCTGT TGCCTGTGGA CAGTTACCAG TCTTGGGCTC TCGCTGGTGG 660
CTGGAACAGC CAGATGTGTT GCCAGGGGAGA ACAGAACCCA CCAGGTCCCT TTTGGAAGGC 720
AGCATTTGCA GACTCCAGCG GGCAGCACCC TCCTGACGCT GGGTGCCTTTC GTCGCGGCGG 780
CAAGAAACGC ATTCCGTACA GCAAGGGGCA GTTGCGGGAG CTGGAGCGGG AGTATGCGGC 840
TAACAAGTTC ATCACCAAGG ACAAGAGGCG CAAGATCTCG GCAGCCACCA GCCTCTCGGA 900
GCGCCAGATT ACCATCTGGT TTCAGAACCG CCGGGTCAAA GAGAAGAAGG TTCTCGCCAA 960
GGTGAAGAAC AGCGCTACCC CT<u>TAA</u>GAGAT CTCCTTGCCT GGGTGGGAGG AGCGAAAGTG 1020
GGGGTGTCCT GGGGAGACCA GAAACCTGCC AAGCCCAGGC TGGGGCCAAG GACTCTGCTG 1080
AGAGGCCCCT AGAGACAACA CCCTTCCCAG GCCACTGGCT GCTGGACTGT CCTCAGGAG 1140
CGGCCTGGGT ACCCAGTATG TGCAGGGAGA CGGAACCCCA TGTGACAGGC CCACTCCACC 1200
AGGGTTCCCA AAGAACCTGG CCCAGTCATA ATCATTCATC CTCACAGTGG CAATAATCAC 1260
GATAACCAGT

45

SEQ ID NO:116 <u>PFJ5 Protein sequence:</u>
Protein Accession #:   NP_006352.1

50
```
1    11    21    31    41    51
|    |     |     |     |     |
|    |     |     |     |     |
```
MEPGNYATLD GAKDIEGLLG AGGGRNLVAH SPLTSHPAAP TLMPAVNYAP LDLPGSAEPP 60
KQCHPCPGVP QGTSPAPVPY GYFGGGYYSC RVSRSSLKPC AQAATLAAYP AETPTAGEEY 120
PSRPTEFAFY PGYPGTYHAM ASYLDVSVVQ TLGAPGEPRH DSLLPVDSYQ SWALAGGWNS 180
QMCCQGEQNP PGPFWKAAFA DSSGQHPPDA CAFRRGRKKR IPYSKGQLRE LEREYAANKF 240
ITKDKRRKIS AATSLSERQI TIWFQNRRVK EKKVLAKVKN SATP

60

<div align="center">SEQ ID NO:117 PFJ4 DNA SEQUENCE</div>

Nucleic Acid Accession #:   NM_005628
Coding sequence: 591-2216   (underlined sequences correspond to start and stop codons)

65
```
1    11    21    31    41    51
|    |     |     |     |     |
|    |     |     |     |     |
```
GTAACCGCTA CTCCCGGACA CCAGACCACC GCCTTCCGTA CACAGGGGCC CGCATCCCAC 60
CCTCCCGGAC CTAAGAGCCT GGGTCCCCTG TTTCCGGAGG TCCGCTTCCC GGCCCCCAGA 120
TTCTGGCATC CCAGCCCTCA GTGTCCAAGA CCCAGGCAGC CCGGGTCCCC GCCTCCCGGA 180
TCCAGGCGTC CGGGATCTGC GCCACCAGAA CCTAGCCTCC TGCAGACCTC CGCCATCTGG 240
70
GGGCACTCAA CCTCCTGGAG CCAAGGGCCC CACGTCCCAC CCAGAGAAAC TCTCGTATTC 300
CCAGCTCCTA GGGCCAAGGA ACCCGGGCGC TCCGAACTCC CAGCTTTCGG ACATCTGGCA 360
CACGGGCAG AGCAGAGAAG CTCAGCGCCC AGCCTGGGGA ATTTAAACAC TCCAGCTTCC 420
AAGAGCCAAG GAACTTCAGT GCTGTGAACT CACAACTCTA AGGAGCCCTC CAAAGTTCCA 480
GTCTCCAGGT GCTGTTACTC AACTCAGTCC TAGGAACGTC GGGTCCTGGG AAGGAGCCCA 540
AGCGCTCCCA GCCAGCTTCC AGGCGCTAAG AAACCCCGGT GCTTCCCATC <u>ATG</u>GTGGCCG 600
ATCCTCCTCG AGACTCCAAG GGGCTCGCAG CGGCGGAGCC CACCGCCAAC GGGGGCCTGG 660
CGCTGGCCTC CATCGAGGAC CAAGGCGCGG CAGCAGGCGG CTACTGCGGT TCCCGGGACC 720
AGGTGCGCCG CTGCCTTCGA GCCAACCTGC TTGTGCTGCT GACAGTGGTG GCCGTGGTGG 780
CCGGCGTGGC GCTGGGACTG GGGGTGTCGG GGGCCGGGGG TGCGCTGGCC TTGGGCCCGG 840

343

AGCGCTTGAG CGCCTTCGTC TTCCCGGGCG AGCTGCTGCT GCGTCTGCTG CGGATGATCA 900
TCTTGCCGCT GGTGGTGTGC AGCTTGATCG GCGGCGCCGC CAGCCTGGAC CCCGGCGCGC 960
TCGGCCGTCT GGGCGCCTGG GCGCTGCTCT TTTTCCTGGT CACCACGCTG CTGGCGTCGG 1020
CGCTCGGAGT GGGCTTGGCG CTGGCTCTGC AGCCGGGCGC CGCCTCCGCC GCCATCAACG 1080
CCTCCGTGGG AGCCGCGGGC AGTGCCGAAA ATGCCCCCAG CAAGGAGGTG CTCGATTCGT 1140
TCCTGGATCT TGCGAGAAAT ATCTTCCCTT CCAACCTGGT GTCAGCAGCC TTTCGCTCAT 1200
ACTCTACCAC CTATGAAGAG AGGAATATCA CCGGAACCAG GGTGAAGGTG CCCGTGGGGC 1260
AGGAGGTGGA GGGGATGAAC ATCCTGGGCT TGGTAGTGTT TGCCATCGTC TTTGGTGTGG 1320
CGCTGCGGAA GCTGGGGCCT GAAGGGGAGC TGCTTATCCG CTTCTTCAAC TCCTTCAATG 1380
AGGCCACCAT GGTTCTGGTC TCCTGGATCA TGTGGTACGC CCCTGTGGGC ATCATGTTCC 1440
TGGTGGCTGG CAAGATCGTG GAGATGGAGG ATGTGGGTTT ACTCTTTGCC CGCCTTGGCA 1500
AGTACATTCT GTGCTGCCTG CTGGGTCACG CCATCCATGG GCTCCTGGTA CTGCCCCTCA 1560
TCTACTTCCT CTTCACCCGC AAAAACCCCT ACCGCTTCCT GTGGGGCATC GTGACGCCGC 1620
TGGCCACTGC CTTTGGGACC TCTTCCAGTT CCGCCACGCT CCGCCTGATG ATGAAGTGCG 1680
TGGAGGAGAA TAATGGCGTG GCCAAGCACA TCAGCCGTTT CATCCTGCCC ATCGGCGCCA 1740
CCGTCAACAT GGACGGTGCC GCGCTCTTCC AGTGCGTGGC CGCAGTGTTC ATTGCACAGC 1800
TCAGCCAGCA GTCCTTGGAC TTCGTAAAGA TCATCACCAT CCTGGTCACG GCCACAGCGT 1860
CCAGCGTGGG GGCAGCGGGC ATCCCTGCTG GAGGTGTCCT CACTCTGGCC ATCATCCTCG 1920
AAGCAGTCAA CCTCCCGGTC GACCATATCT CCTTGATCCT GGCTGTGGAC TGGCTAGTCG 1980
ACCGGTCCTG TACCGTCCTC AATGTAGAAG GTGACGCTCT GGGGGCAGGA CTCCTCCAAA 2040
ATTATGTGGA CCGTACGGAG TCGAGAAGCA CAGAGCCTGA GTTGATACAA GTGAAGAGTG 2100
AGCTGCCCCT GGATCCGCTG CCAGTCCCCA CTGAGGAAGG AAACCCCCTC CTCAAACACT 2160
ATCGGGGGCC CGCAGGGGAT GCCACGGTCG CCTCTGAGAA GGAATCAGTC ATGTAAACCC 2220
CGGGAGGGAC CTTCCCTGCC CTGCTGGGGG TGCTCTTTGG ACACTGGATT ATGAGGAATG 2280
GATAAATGGA TGAGCTAGGG CTCTGGGGGT CTGCCTGCAC ACTCTGGGGA GCCAGGGGCC 2340
CCAGCACCCT CCAGGACAGG AGATCTGGGA TGCCTGGCTG CTGGAGTACA TGTGTTCACA 2400
AGGGTTACTC CTCAAAACCC CCAGTTCTCA CTCATGTCCC CAACTCAAGG CTAGAAAACA 2460
GCAAGATGGA GAAATAATGT TCTGCTGCGT CCCCACCGTG ACCTGCCTGG CCTCCCCTGT 2520
CTCAGGGAGC AGGTCACAGG TCACCATGGG GAATTCTAGC CCCCACTGGG GGGATGTTAC 2580
AACACCATGC TGGTTATTTT GGCGGCTGTA GTTGTGGGGG GATGTGTGTG TGCACGTGTG 2640
TGTGTGTGTG TGTGTGTGTG TGTGTGTGTG TTCTGTGACC TCCTGTCCCC ATGGTACGTC 2700
CCACCCTGTC CCCAGATCCC CTATTCCCTC CACAATAACA GAAACACTCC CAGGGACTCT 2760
GGGGAGAGGC TGAGGACAAA TACCTGCTGT CACTCCAGAG GACATTTTTT TTAGCAATAA 2820
AATTGAGTGT CAACTATTTA AAAAAAAAAA AAAAAA

SEQ ID NO:118 PFJ4 Protein sequence:
Protein Accession #:   NP_005619.1

```
1      11     21     31     41     51
|      |      |      |      |      |
MVADPPRDSK GLAAAEPTAN GGLALASIED QGAAAGGYCG SRDQVRRCLR ANLLVLLTVV   60
AVVAGVALGL GVSGAGGALA LGPERLSAFV FPGELLLRLL RMIILPLVVC SLIGGAASLD  120
PGALGRLGAW ALLFFLVTTL LASALGVGLA LALQPGAASA AINASVGAAG SAENAPSKEV  180
LDSFLDLARN IFPSNLVSAA FRSYSTTYEE RNITGTRVKV PVGQEVEGMN ILGLVVFAIV  240
FGVALRKLGP EGELLIRFFN SFNEATMVLV SWIMWYAPVG IMFLVAGKIV EMEDVGLLFA  300
RLGKYILCCL LGHAIHGLLV LPLIYFLFTR KNPYRFLWGI VTPLATAFGT SSSSATLPLM  360
MKCVEENNGV AKHISRFILP IGATVNMDGA ALFQCVAAVF IAQLSQQSLD FVKIITILVT  420
ATASSVGAAG IPAGGVLTLA IILEAVNLPV DHISLILAVD WLVDRSCTVL NVEGDALGAG  480
LLQNYVDRTE SRSTEPELIQ VKSELPLDPL PVPTEEGNPL LKHYRGPAGD ATVASEKESV  540
M
```

SEQ ID NO:119 PFJ3 DNA SEQUENCE
Nucleic Acid Accession #:   NM_006708
Coding sequence: 88-642   (underlined sequences correspond to start and stop codons)

```
1      11     21     31     41     51
|      |      |      |      |      |
CTAGTTAAGG CGGCACAGGG CCGAGGCGTA GTGTGGGTGA CTCCTCCGTT CCTTGGGTCC   60
CGTCGTCTGT GATACTGCAG TTCAGCCATG GCAGAACCGC AGCCCCCGTC CGGCGGGCCTC  120
ACGGACGAGG CCGCCCTCAG TTGCTGCTCC GACGCGGACC CCAGTACCAA GGATTTTCTA  180
TTGCAGCAGA CCATGCTACG AGTGAAGGAT CCTAAGAAGT CACTGGATTT TTATACTAGA  240
GTTCTTGGAA TGACGCTAAT CCAAAAAATGT GATTTTCCCA TTATGAAGTT TTCACTCTAC  300
TTCTTGGCTT ATGAGGATAA AAATGACATC CCTAAAGAAA AAGATGAAAA AATAGCCTGG  360
GCGCTCTCCA GAAAAGCTAC ACTTGAGCTG ACACACAATT GGGGCACTGA AGATGATGCG  420
ACCCAGAGTT ACCACAATGG CAATTCAGAC CCTCGAGGAT TCGGTCATAT GGAATTGCT  480
GTTCCTGATG TATACAGTGC TTGTAAAAGG TTTGAAGAAC TGGGAGTCAA ATTTGTGAAG  540
AAACCTGATG ATGGTAAAAT GAAAGGCCTG GCATTTATTC AAGATCCTGA TGGCTACTGG  600
ATTGAAATTT TGAATCCTAA CAAAATGGCA ACCTTAATGT AGTGCTGTGA GAATTCTCCT  660
TTGAGATTTC AGAAGAAAGG AAACAATGTG ATTCAAGATA TTTACATACC AGAAGCATCT  720
AGGACTGATG GATCACTGTC CCGATTCAAA TTATTCTTCA GTCCATTTCC CCTTCCTATT  780
TCAGCTGTTC CTTTTCACCT AACTGTTCAG TCATTCTGGT TTTCAAGCAG TGCTTTATCT  840
CATGTCCTTG AATATAGTTG TGTAACTTTA TTTTTTTAGGT AATAATTAGA ACAGTTCCCT  900
TCAGAGGCTG CATTTGCCTT CTTCTGCCAC CTAAATATTA CTTCCCTTCA AATCTGCCTT  960
TGAATCATCA TTTTTTAAAAA AAAATTAACA TGTTTTTTGTT GTAGTTATCT TCTGGGGTTT 1020
CAATTCCTCA GAAACAACTT TTTTCACAAC GGAAAGGAAA GAACACTAGT GTTCTTTCAG 1080
TAAAGTACAA AGTGTTTATT TTACAAAAGA GTAGGTACTC TTGAGAGCAA TTCAAATCAT 1140
```

344

```
GCTGACAAGG ATACTGATAG AAAAAGTGAT TTCTTCTTAT TATAAAGTAC ATTTAAAGTT 1200
CAAGGACTAA CCTTATTTAT TTGGGAAAGG GGAGGAGGAA GGAAATGATA TGGTACCCAG 1260
ACACTGGGCT AGGCTGCAAC TTTATCTCAT TTAATACTCC CAGCTGTCAT GTGAGAAAGA 1320
AAGCAGGCTA GGCATGTGAA ATCACTTTCA TGGATTATTA ATGGATTTAA GAGGGCATCA 1380
ATCAGCTCAA CTCAAGATTT CATAATCATT TTTAGTATTT AGATTGTGCC TCAAAGTTGT 1440
AGTACCTCAC AATACCTCCA CTGGTTTCCT GTTGTAAAAA CCTTCAGTGA GTTTGACCAT 1500
TGTGCTCTTG GCTCTTGGGC TGGAGTACCG TGGTGAGGGA GTAAACACTA GAAGTCTTTA 1560
GTACAAAACT GCTCTAGGGA CACCTGGTGA TTCCTACACA AGTGATGTTT ATATTTCTCA 1620
TAAAGAGTCT TCCCTATCCC AAGGTCTTCA TGATGCCAGT AGCCATATAT GATAAATTAT 1680
GTTCAGTGAT AACTTAGTTA TCAGAAATCA GCTCAGTGGT CTTCCCCGCC ATGATTCACA 1740
TTTGATGAGT TTTTAAAAAT CAAAGTGATT TTGAAAATCT CTAATGGCTC AGAAAATAAA 1800
AACATCCAGT TTGTGGATGA CTATATTTAG ATTTCTCTAG ACTCTAGTGG AAGACCTTTG 1860
GAAAGGCCAT GCCAACCGTG CTTGTACTGC TAGAAGCACT TTATGTTTCC TTTTTGGGTG 1920
AAATGGATTT ATGTGAGTGC TTTAAACAAA TAGCAATACT TATAGACTGA AATAAAATGA 1980
AACTTCAAAT AAG
```

```
1     11    21    31    41    51
|     |     |     |     |     |
MAEPQPPSGG LTDEAALSCC SDADPSTKDF LLQQTMLRVK DPKKSLDFYT RVLGMTLIQK   60
CDFPIMKFSL YFLAYEDKND IPKEKDEKIA WALSRKATLE LTHNWGTEDD ATQSYHNGNS  120
DPRGFGHIGI AVPDVYSACK RFEELGVKFV KKPDDGKMKG LAFIQDPDGY WIEILNPNKM  180
ATLM
```

```
1     11    21    31    41    51
|     |     |     |     |     |
CCGACGCCAG GTCCTGCCGT CCCGCCGACC GTCCGGGAGC GAACCCGTCG TCCCGCACTG   60
GAGTCCGCGA TGGCTTCAGT GACAGATGGT AAACATGGAG TCAAAGATGC CTCTGACCAG  120
AATTTTGACT ACATGTTTAA ACTGCTTATC ATTGGCAACA GCAGTGTTGG CAAGACCTCC  180
TTCCTCTTGC GCTATGCTGA TGACACGTTC ACCCCAGCCT TCGTTAGCAC CGTGGGCATC  240
GACTTCAAGG TGAAGACAGT CTACCGTCAC GAGAAGCGGG TGAAACTGCA GATCTGGGAC  300
ACAGCTGGGC AGGAGCGGTA CCGGACCATC ACAACAGCCT ATTACCGTGG GGCCATGGGC  360
TTCATTCTGA TGTATGACAT CACCAATGAA GAGTCCTTCA ATGCTGTCCA AGACTGGGCT  420
ACTCAGATCA AGACCTACTC CTGGGACAAT GCACAAGTTA TTCTGGTGGG GAACAAGTGT  480
GACATGGAGG AAGAGAGGGT TGTTCCCACT GAGAAGGGCC AGCTCCTTGC AGAGCAGCTT  540
GGGTTTGATT TCTTTGAAGC CAGTGCAAAG GAGAACATCA GTGTAAGGCA GGCCTTTGAG  600
CGCCTGGTGG ATGCCATTTG TGACAAGATG TCTGATTCGC TGGACACAGA CCCGTCGATG  660
CTGGGCTCCT CCAAGAACAC GCGTCTCTCG GACACCCCAC CGCTGCTGCA GCAGAACTGC  720
TCATGCTAGC AAGGCCCACC TTCCTGACCT CCCCTCATTG TGGCCCCACA CCCAAGTCTG  780
CTTCTCCCTG TTACACACTG TCCGCTCT
```

```
1     11    21    31    41    51
|     |     |     |     |     |
MASVTDGKHG VKDASDQNFD YMFKLLIIGN SSVGKTSFLL RYADDTFTPA FVSTVGIDFK   60
VKTVYRHEKR VKLQIWDTAG QERYRTITTA YYRGAMGFIL MYDITNEESF NAVQDWATQI  120
KTYSWDNAQV ILVGNKCDME EERVVPTEKG QLLAEQLGFD FFEASAKENI SVRQAFERLV  180
DAICDKMSDS LDTDPSMLGS SKNTRLSDTP PLLQQNCSC
```

```
1     11    21    31    41    51
|     |     |     |     |     |
ACGCAGAGCG CTGCTGGGCT GCCGGGTCTC CCGCTTCCTC CTCCTGCTCC AAGGGCCTCC   60
TGCATGAGGG CGCGGTAGAG ACCCGGACCC GCGCCGTGCT CCTGCCGTTT CGCTGCGCTC  120
CGCCCGGGCC CGGCTCAGCC AGGCCCCGCG GTGAGCCATG ATTCGCCTCG GGCTCCCCA  180
GTCGCTGGTG CTGCTGACGC TGCTCGTCGC CGCTGTCCTT CGGTGTCAGG GCCAGGATGT  240
CCAGGAGGCT GGCAGCTGTG TGCAGGATGG GCAGAGGTAT AATGATAAGG ATGTGTGGAA  300
GCCGGAGCCC TGCCGGATCT GTGTCTGTGA CACTGGGACT GTCCTCTGCG ACGACATAAT  360
CTGTGAAGAC GTGAAAGACT GCCTCAGCCC TGAGATCCCC TTCGGAGAGT GCTGCCCCAT  420
CTGCCCAACT GACCTCGCCA CTGCCAGTGG GCAACCAGGA CCAAAGGGAC AGAAAGGAGA  480
ACCTGGAGAC ATCAAGGATA TTGTAGGACC CAAAGGACCT CCTGGGCCTC AGGGACCTGC  540
```

```
        AGGGGAACAA GGACCCAGAG GGGATCGTGG TGACAAAGGT GAAAAAGGTG CCCCTGGACC  600
        TCGTGGCAGA GATGGAGAAC CTGGGACCCC TGGAAATCCT GGCCCCCCTG GTCCTCCCGG  660
        CCCCCCTGGT CCCCCTGGTC TTGGTGGAAA CTTTGCTGCC CAGATGGCTG GAGGATTTGA  720
    5   TGAAAAGGCT GGTGGCGCCC AGTTGGGAGT AATGCAAGGA CCAATGGGCC CCATGGGACC  780
        TCGAGGACCT CCAGGCCCTG CAGGTGCTCC TGGGCCTCAA GGATTTCAAG GCAATCCTGG  840
        TGAACCTGGT GAACCTGGTG TCTCTGGTCC CATGGGTCCC CGTGGTCCTC CTGGTCCCCC  900
        TGGAAAGCCT GGTGATGATG GTGAAGCTGG AAAACCTGGA AAAGCTGGTG AAAGGGGTCC  960
        GCCTGGTCCT CAGGGTGCTC GTGGTTTCCC AGGAACCCCA GGCCTTCCTG GTGTCAAAGG 1020
        TCACAGAGGT TATCCAGGCC TGGACGGTGC TAAGGGAGAG GCGGGTGCTC CTGGTGTGAA 1080
   10   GGGTGAGAGT GGTTCCCCGG GTGAGAACGG ATCTCCGGGC CCAATGGGTC TCTCGTGGCCT 1140
        GCCTGGTGAA AGAGGACGGA CTGGCCCTGC TGGCGCTGCG GGTGCCCGAG GCAACGATGG 1200
        TCAGCCAGGC CCCGCAGGTC CTCCGGGTCC TGTCGGTCCT GCTGGTGGTC CTGGCTTCCC 1260
        TGGTGCTCCT GGAGCCAAGG GTGAAGCCGG CCCCACTGGT GCCCGTGGTC CTGAAGGTGC 1320
        TCAAGGTCCT CGCGGTGAAC CTGGTACTCC TGGGTCCCCT GGGCCTGCTG GTGCCTCCGG 1380
   15   TAACCCTGGA ACAGATGGAA TTCCTGGAGC CAAAGGATCT GCTGGTGCTC CTGGCATTGC 1440
        TGGTGCTCCT GGCTTCCCTG GGCCACGGGG TCCTCCTGGC CCTCAAGGTG CAACTGGTCC 1500
        TCTGGGCCCG AAAGGTCAGA CGGGTGAACC TGGTATTGCT GGCTTCAAAG GTGAACAAGG 1560
        CCCCAAGGGA GAACCTGGCC CTGCTGGCCC CCAGGGAGCC CCTGGACCCG CTGGTGAAGA 1620
        AGGCAAGAGA GGTGCCCGTG GAGAGCCTGG TGGCGTTGGG CCCATCGGTC CCCCTGGAGA 1680
   20   AAGAGGTGCT CCCGGAAACC GCGGTTTCCC AGGTCAAGAT GGTCTGGCAG GTCCCAAGGG 1740
        AGCCCCTGGA GAGCGAGGGC CCAGTGGTCT TGCTGGCCCC AAGGGAGCCA ACGGTGACCC 1800
        TGGCCGTCCT GGAGAACCTG GCCTTCCTGG AGCCCGGGGT CTCACTGGCC GCCCTGGTGA 1860
        TGCTGGTCCT CAAGGCAAAG TTGGCCCTTC TGGAGCCCCT GGTGAAGATG GTCGTCCTGG 1920
        ACCTCCAGGT CCTCAGGGGG CTCGTGGGCA GCCTGGTGTC ATGGGTTTCC CTGGCCCCAA 1980
   25   AGGTGCCAAC GGTGAGCCTG GCAAAGCTGG TGAGAAGGGA CTGCCTGGTG CTCCTGGTCT 2040
        GAGGGGTCTT CCTGGCAAAG ATGGTGAGAC AGGTGCTGCA GGACCCCCTG GCCCTGCTGG 2100
        ACCTGCTGGT GAACGAGGCG AGCAGGGTGC TCCTGGGCCA TCTGGGTTCC AGGGACTTCC 2160
        TGGCCCTCCT GGTCCCCCAG GTGAAGGTGG AAAACCAGGT GACCAGGGTG TTCCCGGTGA 2220
        AGCTGGAGCC CCTGGCCTCG TGGGTCCCAG GGGTGAACGA GGTTTCCCAG GTGAACGTGG 2280
   30   CTCTCCCGGT GCCCAGGGCC TCCAGGGTCC CCGTGGCCTC CCCGGCACTC CTGGCACTGA 2340
        TGGTCCCAAA GGTGCATCTG GCCCAGCAGG CCCCCCTGGC GCACAGGGCC CTCCAGGTCT 2400
        TCAGGGAATG CCTGGCGAGA GGGGAGCAGC TGGTATCGCT GGGCCCAAAG GCGACAGGGG 2460
        TGACGTTGGT GAGAAAGGCC CTGAGGGAGC CCCTGGAAAG GATGGTGGAC GAGGCCTGAC 2520
   35   AGGTCCCATT GGCCCCCCTG GCCCAGCTGG TGCTAACGGC GAGAAGGGAG AAGTTGGACC 2580
        TCCTGGTCCT GCAGGAAGTG CTGGTGCTCG TGGCGCTCCA GGTGAACGTG GAGAGACTGG 2640
        CCCCCCCGGA CCAGCGGGAT TTGCTGGGGC TCCTGGTGCT GATGGCCAGC CTGGGGCCAA 2700
        GGGTGAGCAA GGAGAGGCCG GCCAGAAAGG CGATGCTGGT GCCCCTGGTC CTCAGGGCCC 2760
        CTCTGGAGCA CCTGGGCCTC AGGGTCCTAC TGGAGTGACT GGTCCTAAAG GAGCCCGAGG 2820
   40   TGCCCAAGGC CCCCCGGGAG CCACTGGATT CCCTGGAGCT GCTGGCCGCG TTGGACCCCC 2880
        AGGCTCCAAT GGCAACCCTG GACCCCCTGG TCCCCCTGGT CCTTCTGGAA AAGATGGTCC 2940
        CAAAGGTGCT CGAGGAGACA GCGGCCCCCC TGGCCGAGCT GGTGAACCCG GCCTCCAAGG 3000
        TCCTGCTGGA CCCCCTGGCG AGAAGGGAGA GCCTGGAGCT GCCTTTGCTG GCTTAGGCGC 3060
        AGGTCCACCA GGTCCCCAGG GTCTGGCTGG TCAGAGAGGC ATCGTCGGTC TGCCTGGGCA 3120
   45   ACGTGGTGAG AGAGGATTCC CTGGCTTGCC TGGCCCATCG GGTGAGCCCG GCAAGCAGGG 3180
        TGCTCCTGGA GCATCTGGAG ACAGAGGTCC TCCTGGCCCC GTGGGTCCTC CTGGCCTGAC 3240
        GGGTCCTGCA GGTGAACCCG GACGAGAGGG AAGCCCCGGT GCTGATGGCC CCCCTGGCAG 3300
        AGATGGCGCT GCTGGAGTCA AGGGTGATCG TGGTGAGACT GGTGCTGTGG GAGCTCCTGG 3360
        AGCCCCTGGG CCCCCTGGCT CCCCTGGCCC CGCTGGTCCA ACTGGCAAGC AAGGAGACAG 3420
   50   AGGAGAAGCT GGTGCACAAG GCCCCATGGG ACCCTCAGGA CCAGCTGGAG CCCGGGGAAT 3480
        CCAGGGTCCT CAAGGCCCCA GAGGTGACAA AGGAGAGGCT GGAGAGCCTG GCGAGAGAGG 3540
        CCTGAAGGGA CACCGTGGCT TCACTGGTCT GCAGGGTCTG CCCGGCCCTC CTGGTCCTTC 3600
        TGGAGACCAA GGTGCTTCTG GTCCTGCTGG TCCTTCTGGC CCTAGAGGTC CTCCTGGCCC 3660
   55   CGTCGGTCCC TCTGGCAAAG ATGGTGCTAA TGGAATCCCT GGCCCCATTG GGCCTCCTGG 3720
        TCCCCGTGGA CGATCAGGCG AAACCGGTCC TGCTGGTCCT CCTGGAAATC CTGGGCCCCC 3780
        TGGTCCTCCA GGTCCCCCTG GCCCTGGCAT CGACATGTCC GGCCTTTGCTG GCTTAGGCCC 3840
        GAGAGAGAAG GGCCCCGACC CCCTGCAGTA CATGCGGGCC GACCAGGCAG CCGGTGGCCT 3900
        GAGACAGCAT GACGCCGAGG TGGATGCCAC ACTCAAGTCC CTCAACAACC AGATTGAGAG 3960
   60   CATCCGCAGC CCCGAGGGCT CCCGCAAGAA CCCTGCTCGC ACCTGCAGAG ACCTGAAACT 4020
        CTGCCACCCT GAGTGGAAGA GTGGAGACTA CTGGATTGAC CCAAACCAAG GCTGCACCTT 4080
        GGACGCCATG AAGGTTTTCT GCAACATGGA GACTGGCGAG ACTTGCGTCT ACCCCAATCC 4140
        AGCAAACGTT CCCAAGAAGA ACTGGTGGAG CAGCAAGAGC AAGGAGAAGA AACACATCTG 4200
        GTTTGGAGAA ACCATCAATG GTGGCTTCCA TTTCAGCTAT GGAGATGACA ATCTGGCTCC 4260
   65   CAACACTGCC AACGTCCAGA TGACCTTCCT ACGCCTGCTG TCCACGGAAG GCTCCCAGAA 4320
        CATCACCTAC CACTGCAAGA ACAGCATTGC CTATCTGGAC GAAGCAGCTG GCAACCTCAA 4380
        GAAGGCCCTG CTCATCCAGG GCTCCAATGA CGTGGAGATC CGGGCAGAGG GCAATAGCAG 4440
        GTTCACGTAC ACTGCCCTGA AGGATGGCTG CACGAAACAT ACCGGTAAGT GGGCCAAGAC 4500
   70   TGTTATCGAG TACCGGTCAC AGAAGACCTC ACGCCTCCCC ATCATTGACA TTGCACCCAT 4560
        GGACATAGGA GGGCCCGAGC AGGAATTCGG TGTGGACATA GGGCCGGTCT GCTTCTTGTA 4620
        AAAACCTGAA CCCAGAAACA ACACAATCCG TTGCAAACCT AAAGGACCCA AGTACTTTCC 4680
        AATCTCAGTC ACTCTAGGAC TCTGCACTGA ATGGCTGACC TGACCTGATG TCCATTCATC 4740
        CCACCCTCTC ACAGTTCGGA CTTTTCTCCC CTCTCTTTCT AAGAGACCTG AACTGGGCAG 4800
        ACTGCAAAAT AAAATCTCGG TGTTCTATTT ATTTATTGTC TTCCTGTAAG ACCTTCGGGT 4860
   75   CAAGGCAGAG GCAGGAAACT AACTGGTGTG AGTCAAATGC CCCCTGAGTG ACTGCCCCCA 4920
        GCCCAGGCCA GAAGACCTCC CTTCAGGTGC CGGGCGCAGG AACTGTGTGT GTCCTACACA 4980
        ATGGTGCTAT TCTGTGTCAA ACACCTCTGT ATTTTTTAAA ACATCAATTG ATATTAAAAA 5040
        TGAAAAGATT ATTGGAAAGT
```

SEQ ID NO:124 PFJ1 Protein sequence:

Protein Accession #:   NP_001835.2

```
 1    11    21    31    41    51
 |    |     |     |     |     |
MIRLGAPQSL VLLTLLVAAV LRCQGQDVQE AGSCVQDGQR YNDKDVWKPE PCRICVCDTG   60
TVLCDDIICE DVKDCLSPEI PFGECCPICP TDLATASGQP GPKGQKGEPG DIKDIVGPKG  120
PPGPQGPAGE QGPRGDRGDK GEKGAPGPRG RDGEPGTPGN PGPPGPPGPP GPPGLGGNFA  180
AQMAGGFDEK AGGAQLGVMQ GPMGPMGPRG PPGPAGAPGP QGFQGNPGEP GEPGVSGPMG  240
PRGPPGPPGK PGDDGEAGKP GKAGERGPPG PQGARGFPGT PGLPGVKGHR GYPGLDGAKG  300
EAGAPGVKGE SGSPGENGSP GPMGPRGLPG ERGRTGPAGA AGARGNDGQP GPAGPPGPVG  360
PAGGPGFPGA PGAKGEAGPT GARGPEGAQG PRGEPGTPGS PGPAGASGNP GTDGIPGAKG  420
SAGAPGIAGA PGFPGPRGPP GPQGATGPLG PKGQTGEPGI AGFKGEQGPK GEPGPAGPQG  480
APGPAGEEGK RGARGEPGGV GPIGPPGERG APGNRGFPGQ DGLAGPKGAP GERGPSGLAG  540
PKGANGDPGR PGEPGLPGAR GLTGRPGDAG PQGKVGPSGA PGEDGRPGPP GPQGARGQPG  600
VMGFPGPKGA NGEPGKAGEK GLPGAPGLRG LPGKDGETGA AGPPGPAGPA GERGEQGAPG  660
PSGFQGLPGP PGPPGEGGKP GDQGVPGEAG APGLVGPRGE RGFPGERGSP GAQGLQGPRG  720
LPGTPGTDGP KGASGPAGPP GAQGPPGLQG MPGERGAAGI AGPKGDRGDV GEKGPEGAPG  780
KDGGRGLTGP IGPPGPAGAN GEKGEVGPPG PAGSAGARGA PGERGETGPP GPAGFAGPPG  840
ADGQPGAKGE QGEAGQKGDA GAPGPQGPSG APGPQGPTGV TGPKGARGAQ GPPGATGFPG  900
AAGRVGPPGS NGNPGPPGPP GPSGKDGPKG ARGDSGPPGR AGEPGLQGPA GPPGEKGEPG  960
DDGPSGAEGP PGPQGLAGQR GIVGLPGQRG ERGFPGLPGP SGEPGKQGAP GASGDRGPPG 1020
PVGPPGLTGP AGEPGREGSP GADGPPGRDG AAGVKGDRGE TGAVGAPGAP GPPGSPGPAG 1080
PTGKQGDRGE AGAQGPMGPS GPAGARGIQG PQGPRGDKGE AGEPGERGLK GHRGFTGLQG 1140
LPGPPGPSGD QGASGPAGPS GPRGPPGPVG PSGKDGANGI PGPIGPPGPR GRSGETGPAG 1200
PPGNPGPPGP PGPPGPGIDM SAFAGLGPRE KGPDPLQYMR ADQAAGGLRQ HDAEVDATLK 1260
SLNNQIESIR SPEGSRKNPA RTCRDLKLCH PEWKSGDYWI DPNQGCTLDA MKVFCNMETG 1320
ETCVYPNPAN VPKKNWWSSK SKEKKHIWFG ETINGGFHFS YGDDNLAPNT ANVQMTFLRL 1380
LSTEGSQNIT YHCKNSIAYL DEAAGNLKKA LLIQGSNDVE IRAEGNSRFT YTALKDGCTK 1440
HTGKWGKTVI EYRSQKTSRL PIIDIAPMDI GGPEQEFGVD IGPVCFL
```

SEQ ID NO:125 PFH9 DNA SEQUENCE

Nucleic Acid Accession #:   NM_005084
Coding sequence: 162-1487(underlined sequences correspond to start and stop codons)

```
 1    11    21    31    41    51
 |    |     |     |     |     |
GCTGGTCGGA GGCTCGCAGT GCTGTCGGCG AGAAGCAGTC GGGTTTGGAG CGCTTGGGTC   60
GCGTTGGTGC GCGGTGGAAC GCGCCCAGGG ACCCCAGTTC CCGCGAGCAG CTCCGCGCCG  120
CGCCTGAGAG ACTAAGCTGA AACTGCTGCT CAGCTCCCAA GATGGTGCCA CCCAAATTGC  180
ATGTGCTTTT CTGCCTCTGC GGCTGCCTGG CTGTGGTTTA TCCTTTTGAC TGGCAATACA  240
TAAATCCTGT TGCCCATATG AAATCATCAG CATGGGTCAA CAAAATACAA GTACTGATGG  300
CTGCTGCAAG CTTTGGCCAA ACTAAAATCC CCGGGGGAAA TGGGCCTTAT TCCGTTGGTT  360
GTACAGACTT AATGTTTGAT CACACTAATA AGGGCACCTT CTTGCGTTTA TATTATCCAT  420
CCCAAGATAA TGATCGCCTT GACACCCTTT GGATCCCAAA TAAAGAATAT TTTTGGGGTC  480
TTAGCAAATT TCTTGGAACA CACTGGCTTA TGGGCAACAT TTTGAGGTTA CTCTTTGGTT  540
CAATGACAAC TCCTGCAAAC TGGAATTCCC CTCTGAGGCC TGGTGAAAAA TATCCACTTG  600
TTGTTTTTTC TCATGGTCTT GGGGCATTCA GGACACTTTA TTCTGCTATT GGCATTGACC  660
TGGCATCTCA TGGGTTTATA GTTGCTGCTG TAGAACACAG AGATAGATCT GCATCTGCAA  720
CTTACTATTT CAAGGACCAA TCTGCTGCAG AAATAGGGGA CAAGTCTTGG CTCTACCTTA  780
GAACCCTGAA ACAAGAGGAG GAGACCACATA TACGAAATGA GCAGGTACGG CAAAGAGCAA  840
AAGAATGTTC CCAAGCTCTC AGTCTGATTC TTGACATTGA TCATGGAAAG CCAGTGAAGA  900
ATGCATTAGA TTTAAAGTTT GATATGGAAC AACTGAAGGA CTCTATTGAT AGGGAAAAAA  960
TAGCAGTAAT TGGACATTCT TTTGGTGGAG CAACGGTTAT TCAGACTCTT AGTGAAGATC 1020
AGAGATTCAG ATGTGGTATT GCCCTGGATG CATGGATGTT TCCACTGGGT GATGAAGTAT 1080
ATTCCAGAAT TCCTCAGCCC CTCTTTTTTA TCAACTCTGA ATATTTCCAA TATCCTGCTA 1140
ATATCATAAA AATGAAAAAA TGCTACTCAC CTGATAAAGA AAGAAAGATG ATTACAATCA 1200
GGGGTTCAGT CCACCAGAAT TTTGCTGACT TCACTTTTGC AACTGGCAAA ATAATTGGAC 1260
ACATGCTCAA ATTAAAGGGA GACATAGATT CAAATGTAGC TATTGATCTT AGCAACAAAG 1320
CTTCATTAGC ATTCTTACAA AAGCATTTAG GACTTCATAA AGATTTTGAT CAGTGGGACT 1380
GCTTGATTGA AGGAGATGAT GAGAATCTTA TTCCAGGGAC CAACATTAAC ACAACCAATC 1440
AACACATCAT GTTACAGAAC TCTTCAGGAA TAGAGAAATA CAATTAGGAT TAAAATAGGT 1500
TTTTT
```

SEQ ID NO:126 PFH9 Protein sequence:
Protein Accession #:   NP_005075.1

```
 1    11    21    31    41    51
 |    |     |     |     |     |
MVPPKLHVLF CLCGCLAVVY PFDWQYINPV AHMKSSAWVN KIQVLMAAAS FGQTKIPRGN   60
GPYSVGCTDL MFDHTNKGTF LRLYYPSQDN DRLDTLWIPN KEYFWGLSKF LGTHWLMGNI  120
LRLLFGSMTT PANWNSPLRP GEKYPLVVFS HGLGAFRTLY SAIGIDLASH GFIVAAVEHR  180
DRSASATYYF KDQSAAEIGD KSWLYLRTLK QEEETHIRNE QVRQRAKECS QALSLILDID  240
HGKPVKNALD LKFDMEQLKD SIDREKIAVI GHSFGGATVI QTLSEDQRFR CGIALDAWMF  300
PLGDEVYSRI PQPLFFINSE YFQYPANIIK MKKCYSPDKE RKMITIRGSV HQNFADFTFA  360
TGKIIGHMLK LKGDIDSNVA IDLSNKASLA FLQKHLGLHK DFDQWDCLIE GDDENLIPGT  420
NINTTNQHIM LQNSSGIEKY N
```

SEQ ID NO:127 PFH8 DNA SEQUENCE

Nucleic Acid Accession #: NM_015900
Coding sequence: 32-1402 (underlined sequences correspond to start and stop codons)

```
1    11   21   31   41   51
|    |    |    |    |    |
CACGAGCGGC ACGAGGATTT CCAGCTCAGC GATGCCCCCA GGTCCCTGGG AGAGCTGCTT  60
CTGGGTGGGG GGCCTCATTT TGTGGCTCAG CGTTGGAAGT TCAGGGGATG CACCTCCTAC  120
CCCACAGCCA AAGTGCGCTG ACTTCCAGAG CGCCAACCTT TTTGAAGGCA CCGATCTCAA  180
AGTCCAGTTT CTCCTCTTTG TCCCTTCGAA TCCTAGCTGT GGGCAGCTAG TAGAAGGAAG  240
CAGTGACCTC CAAAACTCTG GGTTCAATGC CACTCTGGGA ACCAAACTAA TTATCCATGG  300
ATTCAGGGTT TTAGGAACAA AGCCTTCCTG GATTGACACA TTTATTAGAA CCCTTCTGCG  360
TGCAACGAAT GCTAATGTGA TTGCCGTGGA CTGGATTTAT GGGTCTACAG GAGTCTACTT  420
CTCAGCTGTG AAAAATGTGA TTAAGTTGAG CCTCGAGATC TCCCTTTTCC TCAATAAACT  480
CCTGGTGCTG GGTGTGTCGG AATCCTCAAT CCACATCATT GGTGTTAGCC TGGGGGCCCA  540
CGTTGGGGGC ATGGTGGGAC AGCTCTTCGG AGGCCAGCTG GGACAGATCA CAGGCCTGGA  600
CCCCGCTGGA CCTGAGTACA CCAGGGCCAG TGTGGAAGAG CGCTTGGATG CTGGAGATGC  660
CCTCTTCGTG GAAGCCATCC ACACAGACAC CGACAATTTG GGTATTCGGA TTCCCGTTGG  720
ACATGTGGAC TACTTCGTCA ACGGAGGCCA AGACCAACCT GGCTGCCCCA CCTTCTTTTA  780
CGCAGGTTAT AGTTATCTGA TCTGTGATCA CATGAGGGCT GTGCACCTCT ACATCAGCGC  840
CCTGGAGAAT TCCTGTCCAC TGATGGCCTT TCCCTGTGCC AGCTACAAGG CCTTCCTTGC  900
TGGACGCTGT CTGGATTGCT TTAACCCCTT TCTGCTTTCC TGCCCAAGGA TAGGACTGGT  960
GGAACAAGGT GGTGTCAAGA TAGAGCCGCT CCCCAAGGAA GTGAAAGTCT ACCTCCTGAC  1020
TACTTCCAGT GCTCCGTACT GCATGCATCA CAGCCTCGTG GAGTTTCACT TGAAGGAACT  1080
GAGAAACAAG GACACCAACA TCGAGGTTAC CTTCCTTAGC AGTAACATCA CCTCTTCATC  1140
TAAGATCACC ATACCTAAGC AGCAACGCTA TGGGAAAGGA ATCATAGCCC ATGCCACCCC  1200
ACAATGCCAG ATAAACCAAG TGAAATTCAA GTTTCAGTCT TCCAACCGAG TTTGGAAAAA  1260
AGACCGGACT ACCATTATTG GGAAGTTCTG CACTGCCCTT TTGCCTGTCA ATGACAGAGA  1320
AAAGATGGTC TGCTTACCTG AACCAGTGAA CTTACAAGCA AGTGTGACTG TTTCCTGTGA  1380
CCTGAAGATA GCCTGTGTGT AGTTTAACCT GGGCAGGACA CATCTCCCTG CATTTTTTTT  1440
TTTTTTTTTT GAGAGAGAGG TGTGATGAGG GATGTGTGTG TGCAGCTTAT TGTAGACCAT  1500
TACTACTAAG GAGAAAAGCA AAGCTCTTTC TTATTTTCCT CATAATCAGC TACCCTGGAG  1560
GGGAGGGAGA ACTCATTTTA CAGAACTTGG TTTCCTTTGC CGATCTTATG TACATACCCA  1620
TTTTAGCTTT CCCATGCATA CTTAACTGCA CTTGCTTTAT CTCCTTGGGC ATTCGTACTT  1680
AGGATTCAAT AGAAACATGT ACAGGGTAAA CAATTTTTTA AAAATAAAAC TTCATGGAGT  1740
AAAAAAAAAA AAAAAAAA
```

SEQ ID NO:128 PFH8 Protein sequence:
Protein Accession #: NP_056984.1

```
1    11   21   31   41   51
|    |    |    |    |    |
MPPGPWESCF WVGGLILWLS VGSSGDAPPT PQPKCADFQS ANLFEGTDLK VQFLLFVPSN  60
PSCGQLVEGS SDLQNSGFNA TLGTKLIIHG FRVLGTKPSW IDTFIRTLLR ATNANVIAVD  120
WIYGSTGVYF SAVKNVIKLS LEISLFLNKL LVLGVSESSI HIIGVSLGAH VGGMVGQLFG  180
GQLGQITGLD PAGPEYTRAS VEERLDAGDA LFVEAIHTDT DNLGIRIPVG HVDYFVNGGQ  240
DQPGCPTFFY AGYSYLICDH MRAVHLYISA LENSCPLMAF PCASYKAFLA GRCLDCFNPF  300
LLSCPRIGLV EQGGVKIEPL PKEVKVYLLT TSSAPYCMHH SLVEFHLKEL RNKDTNIEVT  360
FLSSNITSSS KITIPKQQRY GKGIIAHATP QCQINQVKFK FQSSNRVWKK DRTTIIGKFC  420
TALLPVNDRE KMVCLPEPVN LQASVTVSCD LKIACV
```

SEQ ID NO:129 PFH7 DNA SEQUENCE

Nucleic Acid Accession #: NM_014384
Coding sequence: 89-1336 (underlined sequences correspond to start and stop codons)

```
1    11   21   31   41   51
|    |    |    |    |    |
CGTTGCCGGG TCGCAGGTCC CGCCAGTGCG AGCGCAACGG AGGTCGAAGG CGTTCAGACT  60
CTTAGCTGAA CGCGGAGCTG CGGCGGCTAT GCTGTGGAGC GGCTGCCGGC GTTTCGGGGC  120
GCGCCTCGGC TGCCTGCCCG GCGGTCTCCG GGTCCTCGGT CAGACCGGCC ACCGGAGCTT  180
GACCTCCTGC ATCGACCCTT CCATGGGACT TAATGAAGAG CAGAAAGAAT TCAAAAAGT  240
GGCCTTTGAC TTTGCTGCCC GAGAGATGGC TCCAAATATG GCAGAGTGGG ACCAGAAGGA  300
GCTGTTCCCA GTGGATGTGA TGCGGAAGGC AGCCCAGCTA GGCTTCGGAG GGGTCTACAT  360
ACAAACAGAT GTGGGCGGGT CTGGGCTGTC ACGTCTTGAT ACCTCTGTCA TTTTTGAAGC  420
CTTGGCTACA GGCTGCACCA GCACCACAGC CTATATAAGC ATCCACAACA TGTGTGCCTG  480
GATGATTGAT AGCTTCGGAA ATGAGGAACA GAGGCACAAA TTTTGCCCAC CGCTCTGTAC  540
CATGGAGAAG TTTGCTTCCT ACTGCCTCAC TGAACCAGGA ATGGGAGTG ATGCTGCCTC  600
TCTTCTGACC TCCGCTAAGA AACAGGGAGA TCATTACATC CTCAATGGCT CCAAGGCCTT  660
CATCAGTGGT GCTGGTGAGT CAGACATCTA TGTGGTCATG TGCCGAACAG GAGGACCAGG  720
CCCCAAGGGC ATCTCATGCA TAGTTGTTGA GAAGGGGACC CCTGGCCTCA GCTTTGGCAA  780
GAAGGAGAAA AAGGTGGGGT GGAACTCCCA GCCAACACGA GCTGTGATCT TCGAAGACTG  840
TGCTGTCCCT GTGGCCAACA GAATTGGGAG CGAGGGGCAG GGCTTCCTCA TTGCCGTGAG  900
```

```
AGGACTGAAC GGAGGGAGGA TCAATATTGC TTCCTGCTCC CTGGGGGCTG CCCACGCCTC  960
TGTCATCCTC ACCCGAGACC ACCTCAATGT CCGGAAGCAG TTTGGAGAGC CTCTGGCCAG  1020
TAACCAGTAC TTGCAATTCA CACTGGCTGA TATGGCAACA AGGCTGGTGG CCGCGCGGCT  1080
GATGGTCCGC AATGCAGCAG TGGCTCTGCA GGAGGAGAGG AAGGATGCAG TGGCCTTGTG  1140
CTCCATGGCC AAGCTCTTTG CTACAGATGA ATGCTTTGCC ATCTGCAACC AGGCCTTGCA  1200
GATGCACGGG GGCTACGGCT ACCTGAAGGA TTACGCTGTT CAGCAGTACG TGCGGGACTC  1260
CAGGGTCCAC CAGATTCTAG AAGGTAGCAA TGAAGTGATG AGGATACTGA TCTCTAGAAG  1320
CCTGCTTCAG GAGTAGAACC CACACTTGTT CTGGCCTGGT GTTCAGTGCG ACTGCAGTCA  1380
GTGTTGAGTG GTGCCATGTG GGCCGCTCTA TTCCAAAGGA ATCATGGATT AGACCCAAGG  1440
GCTGAGCTCC TCTAGGGCAG GACCTGCACC CTGTGTGTTG GCACCAGCAT CGGGTCTTGG  1500
ACTGGGGCAG AATCCCCAGT GGAACCGGAA GAGCTGGCAT GATGAGAAAC ATCAGAAGAA  1560
CACATACTAC CTTGTTTTCC TAATGCCAGA AGGGTGACCA GTGAAGATTC ACCGTCAAAC  1620
CATGAAAGTC CTTTCTTGGA TCCACTTTAT CTTGATTAGT CTGCATTTTA CTAGTTCACT  1680
GGATCCCTCC TCTAGGGGCC TGGGGACTTT CACTGATGCT CTTCCTGATT CTAGAGCAAA  1740
GGTGTGGGAA GGGGAAATGG AGGAATGCCC TCCTGTCTGT GTCGTTCTCT GTGCCACAGC  1800
TACAGATGCA GAAGGTTTCT CTGGATAGCA CACCTCTGAA TGTAAATCAT GATAAAATGG  1860
ATATTTGGAA ACTTACTCCT AAGCTGTGAT GTAGGGTGTA TTTCTACTTC TGGACTGCCT  1920
CAATATCAAG GGCTGAGACT TTTGAATGTT GAATATTCGT TGGGTTTCAT GTTAAGACGC  1980
CTGTGGTCCA GGAGTGCTAT TCAGTGTTTC TGTTCCTGAT AAACACTTTG AATATTTTTT  2040
TGTGTTTTTG TTTCCTTTTC TGAAGCTGTT CCTCCTTTTA AATATTTTTA ATCACATTGA  2100
TAAAATCTAT CCTTCATCCA CCTCTGGTTC TACTATAGTT GATTTTTATT TTAAATGTTT  2160
AATTGTATTT GATTAAACAC TTAACTGGAT TTTGGAATAA TAAAACTCTC GTCCAATTTG  2220
GCTTTTAAAA AAAAAAAA
```

```
1    11    21    31    41    51
|     |     |     |     |     |
MLWSGCRRFG ARLGCLPGGL RVLVQTGHRS LTSCIDPSMG LNEEQKEFQK VAFDFAAREM  60
APNMAEWDQK ELFPVDVMRK AAQLGFGGVY IQTDVGGSGL SRLDTSVIFE ALATGCTSTT  120
AYISIHNMCA WMIDSFGNEE QRHKFCPPLC TMEKFASYCL TEPGSGSDAA SLLTSAKKQG  180
DHYILNGSKA FISGAGESDI YVVMCRTGGP GPKGISCIVV EKGTPGLSFG KKEKKVGWNS  240
QPTRAVIFED CAVPVANRIG SEGQGFLIAV RGLNGGRINI ASCSLGAAHA SVILTRDHLN  300
VRKQFGEPLA SNQYLQFTLA DMATRLVAAR LMVRNAAVAL QEERKDAVAL CSMAKLFATD  360
ECFAICNQAL QMHGGYGYLK DYAVQQYVRD SRVHQILEGS NEVMRILISR SLLQE
```

```
1    11    21    31    41    51
|     |     |     |     |     |
GCCTGCAGAG AGAGGCACTT TGCACCACAG ACAGATAGCA AGAAGGGAAA GACAGAGAGT  60
GAGAAAAAAG AGGAGTCAGT CGCTCCTGGG GAAGGGGAGAG AGTGAGACTG GGAGAAAGAG  120
AAGCACAGAA AGTGTGTGTA AAACGGAGTA AAGAAAGAAA AAAAAAAAAC TACCCTTAAA  180
GCACATTTAA AAAAAAAAAA CTCTGGCAAT TCAAGAAAGA AACAGGCTAC GTTTAAAGAG  240
CATAGAGACA ATGAAAGGCT AAAGAAAATT TTAAAATCTC TGCCACAGTC TCATAGGTGC  300
TTGGAAATGA AAGTAGAACT GCCTGTCTTT AACGGACTCT GAGGAGGGTA ACTGGATTAG  360
GGACGAGTAC GCCAGCTTTT TTTTTTTTTT TTTTTTTTTT TTTAACATCT TAAATCCTGA  420
AAAAAAAAAA AAAAAAAAAA AAAAGGCAGC AGCTCCGAAT TGAATGAATT GATGGGCACA  480
CTCCAACTGC TGGGCTGGAG AGACTGGACT TAGTCTTGCC ATTTCTGCTT CTTTGAAAGA  540
GGAGACAACT TGGGCTTCCT TTTAATTTAG TTTTTTTTCC CCTTCTCCCC CAACCCCCAA  600
CCTTCCCCCT TACCTCCCCC ACCCCCTTTA TCACCACCCC CCTTTTAAAT AAGAGGGTGA  660
AGGGGAACCA GAGCGCACAA GGGAACTGAC TCAGGAGGCA GAGAAGATGG GCATCCTCAG  720
CGTAGACTTG CTGATCACAC TGCAAATTCT GCCAGTTTTT TTCTCCAACT GCCTCTTCCT  780
GGCTCTCTAT GACTCGGTCA TTCTGCTCAA GCACGTGGTG CTGCTGTTGA GCCGCTCCAA  840
GTCCACTCGC GGAGAGTGGC GGCGCATGCT GACCTCAGAG GGACTGCGCT GCGTCTGGAA  900
GAGCTTCCTC CTCGATGCCT ACAAACAGGT GAAATTGGGT GAGGATGCCC CCAATTCCAG  960
TGTGGTGCAT GTCTCCAGTA CAGAAGGAGG TGACAACAGT GGCAATGGTA CCCAGGAGAA  1020
GATAGCTGAG GGAGCCACAT GCCACCTTCT TGACTTTGCC AGCCCTGAGC GCCCACTAGT  1080
GGTCAACTTT GGCTCAGCCA CTTGACCTCC TTTCACGAGC CAGCTGCCAG CCTTCCGCAA  1140
ACTGGTGGAA GAGTTCTCCT CAGTGGCTGA CTTCCTGCTG GTCTACATTG ATGAGGCTCA  1200
TCCATCGAT GGCTGGGCGA TACCGGGGGA CTCCTCTTTG TCTTTTGAGG TGAAGAAGCA  1260
CCAGAACCAG GAAGATCGAT GTGCAGCAGC CCAGCAGCTT CTGGAGCGTT TCTCCTTGCC  1320
GCCCCAGTGC CGAGTTGTGG CTGACCGCAT GGACAATAAC GCCAACATAG CTTACGGGGT  1380
AGCCTTTGAA CGTGTGTGCA TTGTGCAGAG ACAGAAAATT GCTTATCTGG GAGGAAAGGG  1440
CCCCTTCTCC TACAACCTTC AAGAAGTCCG GCATTGGCTG GAGAAGAATT TCAGCAAGAG  1500
ATGAAAGAAA ACTAGATTAG CTGGTTAAAG GTATGATTAT AAGAGAGCTT ATTGTTTTAA  1560
AAAGTTATAT AAAGGCAAGG AAATTAAGAA CTGAATCCAT ATTTCAACAG AGCCCTATTG  1620
GCTTACTGAA AGACAGGAGT TTATCTATCG GAAGAACATG AATCTCTAAC AGCTCCATAC  1680
TTCTTTCACT ACTCAAATGG CATTGGGCTG AGTAAGTAAC CATATCACCT CTCTTCTTAG  1740
TAAAAAGCCC TATGTGAAAA GATCCCAAGA TGGAGAGGAA GAAACGCTAA TTCAGCATGT  1800
GTTCATTCTG CATTGAGAAG GAACTGATAC ATCTGATGCA TGCTTTGAGA CCAGAAGAAA  1860
AGACTTACCT GAATAATTAC TACATTAGGG AAGCTACTGT CTACGTTAAG ATAAAGGGTA  1920
```

TTGCCTTGGC TCTATTTGGC ATGGATGGAG CCCAGTTGGA AAAATTCCCAA ATATTACAAC 1980
AAGTCCTTGA ACCCAGGCCA TGTGGTTAGA CGTTGGTGTT AAGGTTAGAC CTTATGTTAG 2040
AGTCATTTCT GATGTTCCAG CTTCTAGCCA TGTAGTGCTC TCAGTCTTCA TACCCCAGAA 2100
ATTATTGGTA TATTTGTAGA TACCGAGAAT GATCCCTCAG TCTGAGAGGT TAGAATGATC 2160
ATCTGTAATC TGAGGGTTAA TTTCTAGGCA GGTGGAGAGA GTGGTAAAAA AGAAATGAAA 2220
TTGACAAGCT AGGAAAGAGG AGGCAGAAAG ATTTGGAAAA TTCACAGAGT TTCACCCTTA 2280
AGCTGTAGAG AGTGGGTCAC ATTTGTTAGC CACGGAAACA TAGAAACATA CACAAGGCCA 2340
GAAAAAGAAG AAGGAGCTCA ACTAAAAGTG GCATAGAGAA TACACATATA AAAACAATAT 2400
ATTTGTCATA TGCTCCTAGA GAGGAGAAAG GGGTGATTGA AAGAAAAAAA AATACTTAAA 2460
TATTTGTAAT TGTGAGGGGT TTCTTTTGGA AATAATTACT TTTGAACCAT GTATGTGGTA 2520
TGTATATTTT CAGTGGGTTA ATTATACCCC ATGATACCTA TTAAAGGAAA ACCAGTGGGT 2580
CTGGTGGTGC TGGTCTTTTC CTCCCCATTC CTACAATTTC TATGTGGCCC AAGTCATTCC 2640
TAATCTTGGT CTCTATAGCA GTGTTCTCTC TGAATGCTGA GCTGAAGAAA TTATACGTAC 2700
ATACACACAT ACATACATAC ATACAAATAT ATGTATATAT ATTCTCAGCT GCTGCGGGAG 2760
GTAGGTACCA TGGCCATTCA GCACAGCCTT GATTTCCTCC CAAAGTAGGT GAGCTATAGT 2820
GAAGAATAGG TGCAAACAAA CAAGCTTACT TCCATTGCAA AATAGAAGAA GAGGAAGTTA 2880
GAGATAATTC TGATCAATCA TTTTGGAGGC TTTGTTATAA GGCAACCCCC GGTATATCAT 2940
GGAATTTCCA TTGACATTTG AATTTGGACT TGGATCTTCC CTTGGTCCCA TTAGCTGAGG 3000
TTTAGTAATC TAAAGTCCCT ATAGTATATG ATTATAATGC TATTTTAAAA AATATATATA 3060
TAAAATATTT TTTTCTTTTT AAAATAGACA CTATAGTTTT ACCCATAAGT AATATTTAAA 3120
GATTATAGCT CCCAAAAGAA TGGACCAACC ACTTTCGTAT CATAATTTCT TTTTGGTAAA 3180
TATGAGACTA TTATGAAATC ATAGTATATG ATTGTATTTA AAGGTACAAT CAAAGGATCT 3240
TTTGTCCATT CCATTAATAA CTGAATAAAA AATAAATAAA ATGGATAGAA AAAAACTAAA 3300
GTTGAAAATA CATTCTTAAA CTAGTTGTCT GAAATGAGAA AAGAGTGAGA ACTAGGTGTG 3360
CAAGAACCAA ACGTATTTTA TTTTATTTTT TAAATGGGAG CAACATATCA GTCGTGTCAC 3420
CAGCTGGTAT ATTGTGTAAA TATTAAAGCT CCATTGGGAC TGATTTTTCA TGGCAACATC 3480
AGCTTTCTAA TGTTCTAAAT TCTATAAAAA CCACCCACAA AGAAACAAAG CAAATTTCAT 3540
TATCTAATGA GTTGCTGGAA AATCATATTG AGAATAATTA TTTCAGATTC CTCAGTTGTT 3600
AACTTCTACA TTCAAGGGCT TATCTCTGCC CCCATTGATT TTTAACCTCA AAATGGTGTG 3660
AGATTTACTG TGGAACCCTA AAGCAGTAAA ATAAAAAACC TGGTTGCAGC ACATTCACAC 3720
TGTTGTCCTT AAAATTCCCC TTTTTTCTCT ATGTACGATA AAGTAACAGT ATGTCAGATA 3780
AGCCGGTGGG GGGATGAGAT TAGGCTGAGG CAGTGCTAGT CAACTGGGGG AAAAGGATGA 3840
TGGAAAAATC ACCCAGTTGT GCTATATTTT TAAAGAAGGA GGTCGTTTAT GTGTGCAGAC 3900
AATTCTCCCT GAGGTTAGCC CAATGGAGAA ATGAAGCAGA GGAAGGAAAC ATAGAAAGAC 3960
ATGGGCTATC AGGGAGGAAG ATGTTCAATA GAACATGCAA GAATTTCTGG AAGAAAGGCT 4020
GTGGAAGGGC CAATGGAGAA AATGAATGGA CAAAGCTCAG GAAGCCCTAC GCTATGTAGA 4080
ATGTTCTTGG TGTTATCAGG GTTAAGCCCT GTAATTATGT AACCTATTTA TCGCAACATG 4140
AATTTTTATG ATTTCTTGTG ATGTATTCTT TTATGAAATT AACAAGAACT CATTATTTTG 4200
AGGTAGAGGA AAATCAATGC TTTATCTGAT ATGCTGAGAA ATTATTAGAT TGCCAATACT 4260
CATGTGCGTT TCATGTGTTT TATAAGGTTT GTTCCTTTGA AGAATTGTAG TTCTTAGTCC 4320
CACAGGGAAA TGTGTATCTA TTTATATATC ATAGTATAAA TCTATGATAT ATTTATATCA 4380
TATATAAAAG TCTGAGTTCT CTTTCTTAGT CCCTAATCAT GTTTCTCCCA TAGGCTGTGT 4440
TTACATGGAG CTATCGGTTT AGCCTTTTAA GCTTCATTAG CTTGTCTATT ATTGAAATAG 4500
TTTCCAAGAA ATTTTAGATA TTATCATAAC ATCTGGGTCT ACTCAAACAC TTATTGTTTG 4560
AAAGACTTAT GTCTTGGACC TATCAAAAAC TGACTTTATT TATTGCTTAG TGAAAATACT 4620
AGTGGGATCA ACAATGATTT TCTTGAATGG GCATGAATGG AGATGCCCGC ACAGTAATGT 4680
AGAAATGTTT CATACAGCTA TTAAAATGTA ACTGACCTCC TTAGAGGCAG ATTAGTAACT 4740
GTTCCTACTT TGTATAGCTA AGTGACAGTC ACTTAACTTA CATGACTTTC TTTTTTCACA 4800
TTGGGTCTCT GGTCCTGTGT CTTCACCTCA TTTATAGCAC GTCTCCTTGA TTTTTGGTAG 4860
TATCAACTTC CCAGTGATCT GTTCAGTTAA GTTCTTCTCC CGTTAACCAG GAAGTGCTTA 4920
TTCTCTCATC ACAGTGGGAA GAATAGCCTA TTGTCTTTCA TTTTGCCTGA GTGTATTTTA 4980
CTATTTGGGC TCTGAAATAA AAATTATGAA ATATGGTGAG GTCACATGTT GGTGCTGCCT 5040
TGCTGCATAA AATTCTAGGA GGGCAGGTTA GGAGACAGTT ATGTATGGCC TTTCGGGAAA 5100
ATTCAAAGGG TGGGATTACA AGGGTGTTCC TCAGGCATGC CCCTATGGGC CCTATGTAGG 5160
AGCAAGAAGA ATTGACTGAT TTACAGGACT TCTCTTTATG TCAATCTTAA GAGGATGGAT 5220
GAATCTGGAC ATTTGTTCCA CCCGACCTCT GACTGATGGT TTGGAAAATA ACTTTAATTA 5280
GGATCATATG ACCATTGAAA AAGGAAAAAT GTAGACTCTG ACTTCCGTCC CACTGAAGGA 5340
TTAATGAAAA CCTTTACTAG CATTTAGAGC TTTTTCAGAAC ATCCCCACTG TCATGTGTCT 5400
CAGCAGTGGA GACTGCAAGT AAGGCTTTTA ATTTTAGGAG GTTTTTTTTT TTTTTTTTTT 5460
TTCCCCTAAA TGGTATGGCC AAAAGTCAGA GTTAAAATAT ATATAGTTAG ATTCCAACTT 5520
CCTCCTTCAC TCTAAAAATA GAATCCAAAC CCACTCTTCA TATATGCTTC CAGAATGGGG 5580
CTTAAGTACC AATCTCTGCT TTGCAATGGG CACAATCTTG GTCATGTCCT GAGGCTCTCT 5640
AAGAAAAGAG AGGATCTAGG ATGGGAGAGC TAGAAAGTTG CTAACTGGGA AGAACAAGGC 5700
CCTGAGGGGT TGGTCTACCA ATCTGGGAAG ATTTGAAAAC AAACTTCTCG CAACTGAAGG 5760
AAGGCTGAAG GCTGCTGCAA GTCATTGAGT GACTTTAGGA TGAGCAAAAC ATTGGGCCAC 5820
TTCCTAATGC CCTATGTGTA TAGTACCAGA AGCAAGGTCT CAGACTTAAC AGACCCAGCT 5880
CTGTTCCAAG GTGAGTCTGA ACCAATAGAA AGCAAACATG TGCAGATATC CAAACAAGAC 5940
TGCTCATGCA AGTCGGGGCT GGCTACCCGT CTTAGGCAGC AACAGCAGAG CTCCAGGGAG 6000
CTTATTCAAT ATTTACTGAG ACTTCGAAGA CCCAGCAGAT GTTTAATGAA GTCACTATTT 6060
TGGCTCAAAC CCTCCACTTC TCCCCCTCCC CTCAAAAAGC CAACAGGTAA ACACATAAAT 6120
GAAAGAAACC CACAGAAGGG GATGGGAAAT AAAGAAAATT CTCTCAAGAC TTCTCCAGGC 6180
CCATGTCACT GGTCAGCGTG GTTTTTATGT GTATTAGGAT TGGGGGATGT GAAGAAATAA 6240
GTATCCAGTA CTTTATAACC AAAGCAATTA AATGATATTG GGGTAGGGAA TGTTGGCCAG 6300
TTTTGTTTAG TTTTGCCATC ACATTGTCAC CCAGACCTCA CCTAGCCCCA AGTAATCGGG 6360
CGCCCCGAAG AGGGAGACAG AGATGTGCCA GAGTTGACCC AGTGTGCGGA TGATAACTAC 6420
TGACGAAAGA GTCATCGACC TCAGTTAGTG GTTGGATGTA GTCACATTAG TTTGCCTCTC 6480
CCCATCTTTG TCTCCCTGGC AAGGAGAATA TGCGGGACAT GATGCTAAGA GCCCTGGGTA 6540
AATGTGGTGA GAATGCACGC GTGCATATGC TACACATATG TGCTTCTCAG TTGCAGAAAA 6600
TGAACTGCTT TGGGAGATTA TCAGTAGAAA GAGTGTTATC ATATTGGTGC TGAGTGCTAT 6660

```
GTGTGCTTAT ACAATTTGTT CTTGTATTTT AATAAACTTT GAATAAAAGA ATAAAAAAAA 6720
AAAAAAAAAA AAAAA
```

SEQ ID NO:132 PFH6 Protein sequence:
Protein Accession #:   NP_054644.1

```
1     11    21    31    41    51
|     |     |     |     |     |
MGILSVDLLI TLQILPVFFS NCLFLALYDS VILLKHVVLL LSRSKSTRGE WRRMLTSEGL   60
RCVWKSFLLD AYKQVKLGED APNSSVVHVS STEGGDNSGN GTQEKIAEGA TCHLLDFASP   120
ERPLVVNFGS ATXPPFTSQL PAFRKLVEEF SSVADFLLVY IDEAHPSDGW AIPGDSSLSF   180
EVKKHQNQED RCAAAQQLLE RFSLPPQCRV VADRMDNNAN IAYGVAFERV CIVQRQKIAY   240
LGGKGPFSYN LQEVRHWLEK NFSKRXKKTR LAG
```

SEQ ID NO:133 PFH5 DNA SEQUENCE

Nucleic Acid Accession #:   NM_001141
Coding sequence: 72-2102   (underlined sequences correspond to start and stop codons)

```
1     11    21    31    41    51
|     |     |     |     |     |
CAGGCGTGTC CCAGGGGGAG CCCCGCTCTG CAGCCCTGTG CGCCGTAGAG AGCTGGACTT   60
AGGCTGGCAG CATGGCCGAG TTCAGGGTCA GGGTGTCCAC CGGAGAAGCC TTCGGGGCTG   120
GCACATGGGA CAAAGTGTCT GTCAGCATCG TGGGGACCCG GGGAGAGAGC CCCCCACTGC   180
CCCTGGACAA TCTCGGCAAG GAGTTCACTG CGGGCGCTGA GGAGGACTTC AGGGTGACGC   240
TCCCGGAGGA CGTAGGCCGA GTGCTGCTGC TGCGCGTGCA CAAGGCGCCC CCAGTGCTGC   300
CCCTGCTGGG GCCCCTGGCC CCGGATGCCT GGTTCTGCCG CTGGTTCCAG CTGACACCGC   360
CGCGGGGCGG CCACCTCCTC TTCCCCTGCT ACCAGTGGCT GGAGGGGGCG GGGACCCTGG   420
TGCTGCAGGA GGGTACAGCC AAGGTGTCCT GGGCAGCACA CCACCCTGTG CTCCAGCAAC   480
AGCGCCAGGA GGAGCTTCAG GCCCGGCAGG AGATGTACCA GTGGAAGGCT TACAACCCAG   540
GTTGGCCTCA CTGCCTGGAT GAAAAGACAG TGGAAGACTT GGAGCTCAAT ATCAAATACT   600
CCACAGCCAA GAATGCCAAC TTTTATCTAC AAGCTGGCTC TGCTTTTGCA GAGATGAAAA   660
TCAAGGGGTT GCTGGACCGC AAGGGGCTCC TGGAGGAGTCT GGAGTGAGATG AAAAGGATCT   720
TCAACTTCCG GAGGACCCCA GCAGCTGAGC ACGCATTTGA GCACTGGCAG GAGGATGCCT   780
TCTTCGCCTC CCAGTTCCTG AATGGTCTCA ACCCTGTCCT GATCCGCCGC TGTCACTACC   840
TCCCAAAGAA CTTCCCCGTC ACTGATGCCA TGGTGGCCTC ATTGTTGGGT CCTGGGACCA   900
GCTTGCAGGC TGAGCTAGAG AAGGGGCTCCC TGTTCTTGGT GGATCACGGC ATCCTCTCTG   960
GCATCCAGAC CAATGTCATT AATGGGAAGC CGCAGTTCTC TGCGGCCCCA ATGACCCTGC   1020
TATACCAGAG CCCAGGCTGC GGGCCGCTGC TGCCTCTCGC CATCCAGCTC AGCCAGACCC   1080
CCGGCCCAAA CAGCCCCATC TTCCTGCCCA CTGATGACAA GTGGGACTGG TTGCTGGCCA   1140
AGACCTGGGT GCGCAATGCC GAGTTCTCCT TCCATGAGGC CCTCACGCAC CTGCTGCACT   1200
CACATCTGCT GCCTGAGGTC TTCACCCTGG CTACCCTGCG TCAGCTGCCC CACTGCCACC   1260
CTCTCTTCAA GCTGCTGATC CCGCACACCC GATACACCCT GCACATCAAC ACACTCGCCC   1320
GGGAGCTGCT TATCGTGCCA GGGCAGGTGG TGGACAGGTC CACAGGCATC GGCATTGAAG   1380
GCTTCTCTGA GTTGATACAG AGGAACATGA AGCAGCTGAA CTATTCTCTC CTGTGTCTGC   1440
CTGAGGATAT CCGGACCCGA GGAGTTGAAG ACATCCCAGG CTACTACTAC CGTGATGATG   1500
GGATGCAGAT TTGGGGTGCA GTGGAACGCT TTGTCTCTGA AATCATCGGT ATCTACTACC   1560
CAAGTGATGA GTCTGTCCAA GATGACAGAG AGCTCCAGGC CTGGGTCAGA GAGATCTTCT   1620
CCAAGGGCTT CCTAAACCAG GAGAGCTCAG GTATCCCTTC CTCACTGGAG ACCCGGGAAG   1680
CCCTGGTGCA GTATGTCACC ATGGTGATAT TCACCTGCTC AGCCAAGCAT GCGGCTGTCA   1740
GTGCAGGGCA GTTTGACTCC TGTGCTTGGA TGCCCAACCT GCCACCCAGC ATGCAGCTGC   1800
CACCACCCAC CTCCAAAGGC CTGGCAACAT GCGAGGGCTT CATAGCCACC CTCCCACCTG   1860
TCAATGCCAC ATGTGATGTC ATCCTTGCTC TCTGGTTGCT GAGCAAGGAG CCTGGAACCC   1920
AAAGGCCCCT GGGCACCTAT CCGGATGAGC ACTTCACAGA GGAGGCCCCT CGGCGGAGCA   1980
TCGCCACCTT CCAGAGCCGC CTGGCCCAGA TCTCGAGGGG CATCCAGGAG CGGAACCGGG   2040
GCCTGGTGCT GCCCTACACC TACCTAGACC CTCCCCTCAT CGAGAACAGC GTCTCCATCT   2100
AAATCCCAGG GGAACACAGG CCCAGATGAC ATCCCTTTGA CCACATCGCT CTAGGATAAC   2160
TGGCACCCAG AGAAAAGGAC TCCTCAGAAA AAACAGGCCC CCATGTGCCT CTCCTGGGAC   2220
AACCAGACTC TGTAACTCAC CCCCACCACC ATACACACAC ACAAAAACAG AAACAAAATC   2280
AAAACAGAGA AAGCAGAAAA TCTACCAAGA ACAGAGTCTC AGACACAGAAC CACTGAGTCT   2340
TTTGGAGGCT CCAAGCCTCA AAGTGCCCGC AGAGCCCACC TTGAGGGTTT TGCTAGTTGG   2400
TTTTGTTTTG CGTTTACAGC CGTGGGGGGA AGCACATAAT CCCGCCCCAG GCCCACTAG   2460
CATCCACTGA TTGGACCTTA TGGTCACCCA ACTCAAGGAC AGCCACCAAG AAGTGGCTGC   2520
CAAAGAGACT GGGCGCAGTG GCTCATGCCC ATAATCCCAG CACTTTGGGA GATGGAGGCG   2580
GGAAAATCAT TTGAGGTCAG AAGTTCAAGG CCAGCCTGGA CGACATAGCG AGACTCCACC   2640
TCTACCAAAA AATAAAAAATT AAAAAACAAA AAAAAAAAAA AAAAA
```

SEQ ID NO:134 PFH5 Protein sequence:
Protein Accession #:   NP_001132.1

```
1     11    21    31    41    51
|     |     |     |     |     |
MAEFRVRVST GEAFGAGTWD KVSVSIVGTR GESPPLPLDN LGKEFTAGAE EDFQVTLPED   60
VGRVLLLRVH KAPPVLPLLG PLAPDAWFCR WFQLTPPRGG HLLFPCYQWL EGAGTLVLQE   120
GTAKVSWADH HPVLQQQRQE ELQARQEMYQ WKAYNPGWPH CLDEKTVEDL ELNIKYSTAK   180
```

NANFYLQAGS AFAEMKIKGL LDRKGLWRSL NEMKRIFNFR RTPAAEHAFE HWQEDAFFAS 240
QFLNGLNPVL IRRCHYLPKN FPVTDAMVAS LLGPGTSLQA ELEKGSLFLV DHGILSGIQT 300
NVINGKPQFS AAPMTLLYQS PGCGPLLPLA IQLSQTPGPN SPIFLPTDDK WDWLLAKTWV 360
RNAEFSFHEA LTHLLHSHLL PEVFTLATLR QLPHCHPLFK LLIPHTRYTL HINTLARELL 420
IVPGQVVDRS TGIGIEGFSE LIQRNMKQLN YSLLCLPEDI RTRGVEDIPG YYYRDDGMQI 480
WGAVERFVSE IIGIYYPSDE SVQDDRELQA WVREIFSKGF LNQESSGIPS SLETREALVQ 540
YVTMVIFTCS AKHAAVSAGQ FDSCAWMPNL PPSMQLPPPT SKGLATCEGF IATLPPVNAT 600
CDVILALWLL SKEPGDQRPL GTYPDEHFTE EAPRRSIATF QSRLAQISRG IQERNRGLVL 660
PYTYLDPPLI ENSVSI

<div align="center">SEQ ID NO:135 PFH4 DNA SEQUENCE</div>

Nucleic Acid Accession #:   NM_002742
Coding sequence: 236-2974 (underlined sequences correspond to start and stop codons)

```
1      11      21      31      41      51
|       |       |       |       |       |
GAATTCCTTC TCTCCTCCTC CTCGCCCTTC TCCTCGCCCT CCTCCTCCTC CTCGCCCTCC  60
CCTCCCGATC CTCATCCCCT TGCCCTCCCC CAGCCCAGGG ACTTTTCCGG AAAGTTTTTA  120
TTTTCCGTCT GGGCTCTCGG AGAAAGAAGC TCCTGGCTCA GCGGCTGCAA AACTTTCCTG  180
CTGCCGCGCC GCCAGCCCCC GCCCTCCGCT GCCCGGCCCT GCCCCCGCCC GAGCGATGAG  240
CGCCCCTCCG GTCCTGCGGC CGCCCAGTCC GCTGCTGCCC GTGGCGGCGG CAGCTGCCGC  300
AGCGGCCGCC GCACTGGTCC CAGGGTCCGG GCCCGGGCCC GCGCCGTTCT TGGCTCCTGT  360
CGCGGCCCCG GTCGGGGGCA TCTCGTTCCA TCTGCAGATC GGCCTGAGCC GTGAGCCGGT  420
GCTGCTGCTG CAGGACTCGT CCGGGGACTA CAGCCTGGCG CACGTCCGCG AGATGGCTTG  480
CTCCATTGTC GACCAGAAGT TCCCTGAATG TGGTTTCTAC GGAATGTATG ATAAGATCCT  540
GCTTTTTCGC CATGACCCTA CCTCTGAAAA CATCCTTCAG CTGGTGAAAG CGGCCAGTGA  600
TATCCAGGAA GGCGATCTTA TTGAAGTGGT CTTGTCAGGT TCCGCCACCT TTGAAGACTT  660
TCAGATTCGT CCCCACGCTC TCTTTGTTCA TTCATACAGA GCTCCAGCTT TCTGTGATCA  720
CTGTGGAGAA ATGCTGTGGG GGCTGGTACG TCAAGGTCTT AAATGTGAAG GGTGTGGTCT  780
GAATTACCAT AAGAGATGTG CATTTAAAAT ACCCAACAAT GTCAGCGGTG TGAGGCGGAG  840
AAGGCTCTCA AACGTTTCCC TCACTGGGGT CAGCACCATC CGCACATCAT CTGCTGAACT  900
CTCTACAAGT GCCCCTGATG AGCCCCTTCT GCAAAAATCA CCATCAGAGT CGTTTATTGG  960
TCGAGAGAAG AGGTCAAATT CTCAATCATA CATTGGACGA CCAATTCACC TTGACAAGAT  1020
TTTGATGTCT AAAGTTAAAG TGCCGCACAC ATTTGTCATC CACTCCTACA CCCGGCCCAC  1080
AGTGTGCCAG TACTGCAAGA AGCTTCTGAA GGGGCTTTTC AGGCAGGGCT TGCAGTGCAA  1140
AGATTGCAGA TTCAACTGCC ATAAACGTTG TGCACCGAAA GTACCAAACA ACTGCCTTGG  1200
CGAAGTGACC ATTAATGGAG ATTTGCTTAG CCCTGGGGCA GAGTCTGATG TGGTCATGGA  1260
AGAAGGGAGT GATGACAATG ATAGTGAAAG GAACAGTGGG CTCATGGATG ATATGGAAGA  1320
AGCAATGGTC CAAGATGCAG AGATGGCAAT GGCAGAGTGC CAGAACGACA GTGGCGAGAT  1380
GCAAGATCCA GACCCAGACC ACGAGGACGC CAACAGAACC ATCAGTCCAT CAACAAGCAA  1440
CAATATCCCA CTCATGAGGG TAGTGCAGTC TGTCAAACAC ACGAAGAGGA AAAGCAGCAC  1500
AGTCATGAAA GAAGGATGGA TGGTCCACTA CACCAGCAAG GACACGCTGC GGAAACGGCA  1560
CTATTGGAGA TTGGATAGCA AATGTATTAC CCTCTTTCAG AATGACACAG GAAGCAGGTA  1620
CTACAAGGAA ATTCCTTTAT CTGAAATTTT GTCTCTGGAA CCAGTAAAAA CTTCAGCTTT  1680
AATTCCTAAT GGGGCCAATC CTCATTGTTT CGAAATCACT ACGGCAAATG TAGTGTATTA  1740
TGTGGGAGAA AATGTGGTCA ATCCTTCCAG CCCATCACCA AATAACAGTG TTCTCACCAG  1800
TGGCGTTGGT GCAGATGTGG CCAGGATGTG GGAGATAGCC ATCCAGCATG CCCTTATGCC  1860
CGTCATTCCC AAGGGCTCCT CCGTGGGTAC AGGAACCAAC TTGCACAGAG ATATCTCTGT  1920
GAGTATTTCA GTATCAAATT GCCAGATTCA AGAAAATGTG GACATCAGCA CAGTATATCA  1980
GATTTTTCCT GATGAAGTAC TGGGTTCTGG ACAGTTTGGA ATTGTTTATG GAGGAAAACA  2040
TCGTAAAACA GGAAGAGATG TAGCTATTAA AATCATTGAC AAATTACGAT TTCCAACAAA  2100
ACAAGAAAGC CAGCTTCGTA ATGAGGTTGC AATTCTACAG AACCTTCATC ACCCTGGTGT  2160
TGTAAATTTG GAGTGTATGT TTGAGACGCC TGAAAGAGTG TTTGTTGTTA TGGAAAAACT  2220
CCATGGAGAC ATGCTGGAAA TGATCTTGTC AAGTGAAAAG GGCAGGTTGC CAGAGCACAT  2280
AACGAAGTTT TTAATTACTC AGATACTCGT GGCTTTGCGG CACCTTCATT TTAAAAAATAT  2340
CGTTCACTGT GACCTCAAAC CAGAAAATGT GTTGCTAGCC TCAGCTGATC CTTTTCCTCA  2400
GGTGAAACTT TGTGATTTTG GTTTTGCCCG GATCATTGGA GAGAAGTCTT TCCGGAGGTC  2460
AGTGGTGGGT ACCCCCGCTT ACCTGGCTCC TGAGGTCCTA AGGAACAAGG GCTACAATCG  2520
CTCTCTAGAC ATGTGGTCTG TTGGGGTCAT CATCTATGTA AGCCTAAGCG GCACATTCCC  2580
ATTTAATGAA GATGAAGACA TACACGACCA AATTCAGAAT GCAGCTTTCA TGTATCCACC  2640
AAATCCCTGG AAGGAAATAT CTCATGAAGC CATTGATCTT ATCAACAATT TGCTGCAAGT  2700
AAAAATGAGA AAGCGCTACA GTGTGGATAA GACCTTGAGC CACCCTTGGC TACAGGACTA  2760
TCAGACCTGG TTAGATTTGC GAGAGCTGGA ATGCAAAATC GGGGAGCGCT ACATCACCCA  2820
TGAAAGTGAT GACCTGAGGT GGGAGAAGTA TGCAGGCGAG CAGCGGCTGC AGTACCCCAC  2880
ACACCTGATC AATCCAAGTG CTAGCCACAG TGACACTCCT GAGACTGAAG AAACAGAAAT  2940
GAAAGCCCTC GGTGAGCGTG TCAGCATCCT CTGAGTTCCA TCTCCTATAA TCTGTCAAAA  3000
CACTGTGGAA CTAATAAAATA CATACGGTCA GGTTTAACAT TTGCCTTGCA GAACTGCCAT  3060
TATTTTCTGT CAGATGAGAA CAAAGCTGTT AAACTGTTAG CACTGTTGAT GTATCTGAGT  3120
TGCCAAGACA AATCAACAGA AGCATTTGTA TTTTGTGTGA CCAACTGTGT TGTATTAACA  3180
AAAGTTCCCT GAAACACGAA ACTTGTTATT GTGAATGATT CATGTTATAT TTAATGCATT  3240
AAACCTGTCT CCACTGTGCC TTTGCAAATC AGTGTTTTTC TTACTGGAGC TTCATTTTGG  3300
TAAGAGACAG AATGTATCTG TGAAGTAGTT CTGTTTGGTG TGTCCCATTG GTGTTGTCAT  3360
TGTAAACAAA CTCTTGAAGA GTCGATTATT TCCAGTGTTC TATGTAACAAC TCCAAAACCC  3420
ATGTGGGAAA AAAATGAATG AGGAGGGTAG GGAATAAAAT CCTAAGACAC AAATTGCATGA  3480
ACAAGTTTTA ATGTATAGTT TTGAATCCTT TGCCTGCCTG GTGTGCCTCA GTATATTAA  3540
ACTCAAGACA ATGCACCTAG CTGTGCAAGA CCTAGTGCTC TTAAGCCTAA ATGCCTTAGA  3600
AATGTAAACT GCCATATATA ACAGATACAT TTCCCTCTTT CTTATAATAC TCTGTTGTAC  3660
```

<div align="center">352</div>

TATGGAAAAT CAGCTGCTCA GCAACCTTTC ACCTTTGTGT ATTTTTCAAT AATAAAAAAT 3720
ATTCTTGTCA AAAAAAAAAA AA

SEQ ID NO:136 PFH4 Protein sequence:
Protein Accession #:   NP_002733.1

```
 1     11     21     31     41     51
 |     |      |      |      |      |
MSAPPVLRPP SPLLPVAAAA AAAAAALVPG SGPGPAPFLA PVAAPVGGIS FHLQIGLSRE   60
PVLLLQDSSG DYSLAHVREM ACSIVDQKFP ECGFYGMYDK ILLFRHDPTS ENILQLVKAA  120
SDIQEGDLIE VVLSRSATFE DFQIRPHALF VHSYRAPAFC DHCGEMLWGL VRQGLKCEGC  180
GLNYHKRCAF KIPNNCSGVR RRRLSNVSLT GVSTIRTSSA ELSTSAPDEP LLQKSPSESF  240
IGREKRSNSQ SYIGRPIHLD KILMSKVKVP HTFVIHSYTR PTVCQYCKKL LKGLFRQGLQ  300
CKDCRFNCHK RCAPKVPNNC LGEVTINGDL LSPGAESDVV MEEGSDDNDS ERNSGLMDDM  360
EEAMVQDAEM AMAECQNDSG EMQDPDPDHE DANRTISPST SNNIPLMRVV QSVKHTKRKS  420
STVMKEGWMV HYTSKDTLRK RHYWRLDSKC ITLFQNDTGS RYYKEIPLSE ILSLEPVKTS  480
ALIPNGANPH CFEITTANVV YYVGENVVNP SSPSPNNSVL TSGVGADVAR MWEIAIQHAL  540
MPVIPKGSSV GTGTNLHRDI SVSISVSNCQ IQENVDISTV YQIFPDEVLG SGQFGIVYGG  600
KHRKTGRDVA IKIIDKLRFP TKQESQLRNE VAILQNLHHP GVVNLECMFE TPERVFVVME  660
KLHGDMLEMI LSSEKGRLPE HITKFLITQI LVALRHLHFK NIVHCDLKPE NVLLASADPF  720
PQVKLCDFGF ARIIGEKSFR RSVVGTPAYL APEVLRNKGY NRSLDMWSVG VIIYVSLSGT  780
FPFNEDEDIH DQIQNAAFMY PPNPWKEISH EAIDLINNLL QVKMRKRYSV DKTLSHPWLQ  840
DYQTWLDLRE LECKIGERYI THESDDLRWE KYAGEQRLQY PTHLINPSAS HSDTPETEET  900
EMKALGERVS IL
```

SEQ ID NO:137 PFH3 DNA SEQUENCE

Nucleic Acid Accession #:   X95425
Coding sequence: 712-3825 (underlined sequences correspond to start and stop codons)

```
 1     11     21     31     41     51
 |     |      |      |      |      |
AATGGTCAGT CAATACATTA TAACATAATA CACCAAATGC TAGAATAGAA GGGGAGGGGG   60
GCACACATAA TGACTCACTG CTGGAAGAAG GGTGCATCAG TGAATTAAAA AATGTCCCTC  120
CCCTCTTCAG CACTCAGCGC GCAGCTATTT CCTTCTGCCA GTCTCTTTGA ACTCTGGATC  180
TTTGCTTTTG CTCGCTGCTC TCCTGTTTTT CATTCTCCAC ATTTTCTCAA TCCTCTTTCT  240
TTATCCTTAG CCACCCTGCT TTTTTCCTCC TTTTTTAAAA AATCGGAGAT TTCGTCTTAA  300
AATGATTTGT CTTCCTTACC TTCGTCCATT TCAACACTGA AGGCTGCAAA GAACTTCACC  360
TTTCCCCTAG TGGTATTTAA AAATTCTCAA TCCGTAAAAA GTCTTTTTGA AAGGCAAAGG  420
AACAGGACCC AGACCCTCTC GACACCCTTG ATCCGAGTCA GATCTGCACT AGCAACCAGA  480
ACTAATATTT CATTTAACCC ACCAAAAGGG GGAGGCGAGA GGAGCCAGAA GCAAACTTCA  540
TCTGTCTCAG ACGGATCCGT GGTTCCTACA TTTGGAGGAG CCGCGTGTCA GAAGGCGTAG  600
GACCCCAAGG GGGGACAAGG AGGACTCCCG AGTCTCCCTT CTCCGCTCTC CGAGACCGAA  660
GAGGTGGACT GAGCCGCTCG GGACAGCGGC ACCGGAGGAG GCTCGGAGAA GATGCGGGGC  720
TCGGGGCCCC GGGGTGCGGG ACACCGGCGG CCCCCAAGCG GCGGCGGCGA CACCCCCATC  780
ACCCCAGCGT CCCTGGCCGG CTGCTACTCT GCACCTCGAC GGGCTCCCCT CTGGACGTGC  840
CTTCTCCTGT GCGCCGCACT CCGGACCCTC CTGGCCAGCC CCAGCAACGA AGTGAATTTA  900
TTGGATTCAC GCACTGTCAT GGGGGACCTG GGATGGATTG CTTTTTCCAAA AAATGGGTGG  960
GAAGAGATTG GTGAAGTGGA TGAAAATTAT GCCCCTATCC ACACATACCA AGTATGCAAA 1020
GTGATGGAAC AGAATCAGAA TAACTGGCTT TTGACCAGTT GGATCTCCAA TGAAGGTGCT 1080
TCCAGAATCT TCATAGAACT CAAATTTACC CTGCGGGACT GCAACAGCCT TCCTGGAGGA 1140
CTGGGGACCT GTAAGGAAAC CTTTAATATG TATTACTTTG AGTCAGATGA TCAGAATGGG 1200
AGAAACATCA AGGAAAACCA ATACATCAAA ATTGATACCA TTGCTGCCGA TGAAAGCTTT 1260
ACAGAACTTG ATCTTGGTGA CCGTGTTATG AAACTGAATA CAGAGGTCAG AGATGTAGGA 1320
CCTCTAAGCA AAAAGGGATT TTATCTTGCT TTTCAAGATG TTGGTGCTTG CATTGCTCTG 1380
GTTTCTGTGC GTGTATACTA TAAAAAATGC CCTTCTGTGG TACGACACTT GGCTGTCTTC 1440
CCTGACACCA TCACTGGAGC TGATTCTTCC CAATTGCTCG AAGTGTCAGG CTCCTGTGTC 1500
AACCATTCTG TGACCGATGA ACCTCCCAAA ATGCACTGCA GCGCCGAAGG GGAGTGGCTG 1560
GTGCCCATCG GGAAATGCAT GTGCAAGGCA GGATATGAAG AGAAAAATGG CACCTGTCAA 1620
GTGTGCAGAC CTGGGTTCTT CAAAGCCTCA CCTCACATCC AGAGCTGCGG CAAATGTCCA 1680
CCTCACAGTT ATACCCATGA GGAAGCTTCA ACCTCTTGTG TCTGTGAAAA GGATTATTTC 1740
AGGAGAGAGT CTGATCCACC CACAAATGGCA TGCACAAGAC CCCCCTCTGC TCCTCGGAAT 1800
GCCATCTCAA ATGTTAATGA AACTAGTGTC TTTCTGGAAT GGATTCCGCC TGCTGACACT 1860
GGTGGAAGGA AAGACGTGTC ATATTATATT GCATGCAAGA AGTGCAACTC CCATGCAGGT 1920
GTGTGTGAGG AGTGTGGCGG TCATGTCAGG TACCTTCCCC GGCAAAGCGG CCTGAAAAAC 1980
ACCTCTGTCA TGATGGTGGA TCTACTCGCT CACACAAACT ATACCTTTGA GATTGAGGCA 2040
GTGAATGGAG TGTCCGACTT GAGCCCAGGA GCCCGGCAGT ATGTGTCTGT AAATGTAACC 2100
ACAAATCAAG CAGCTCCATC TCCAGTCACC AATGTGAAAA AAGGGAAAAT TGCAAAAAAC 2160
AGCATCTCTT TGTCTTGGCA AGAACCAGAT CGTCCCAATG GAATCATCCT AGAGTATGAA 2220
ATCAAGCATT TTGAAAAGGA CCAAGAGACC AGCTACACGA TTATCAAATC TAAAGAGACA 2280
ACTATTACTG CAGAGGGGCTT GAAACCAGCT TCAGTTTATG TCTTCCAAAT TCGAGCACGT 2340
ACAGCAGCAG GCTATGGTGT CTTCAGTCGA AGATTTGAGT TTGAAACCAC CCCAGTGTTT 2400
GCAGCATCCA GCGATCAAAG CCAGATTCCT GTAATTGCTG TGTCTGTGAC AGTAGGAGTC 2460
ATTTTGTTGG CAGTGGTTAT CGGCGTCCTC CTCAGTGGAA GTTGCTGCGA ATGTGGCTGT 2520
GGGAGGGCTT CTTCCCCTGTG CGCTGTTGCC CATCCAATCC TAATATGGCG GTGTGGCTAC 2580
AGCAAAGCAA AACAAGATCC AGAAGAGGAA AAGATGCATT TTCATAAATGG GCACATTAAA 2640
```

353

CTGCCAGGAG TAAGAACTTA CATTGATCCA CATACCTATG AGGATCCCAA TCAAGCTGTC 2700
CACGAATTTG CCAAGGAGAT AGAAGCATCA TGTATCACCA TTGAGAGAGT TATTGGAGCA 2760
GGTGAATTTG GTGAAGTTTG TAGTGGACGT TTGAAACTAC CAGGAAAAAG AGAATTACCT 2820
GTGGCTATCA AAACCCTTAA AGTAGGCTAT ACTGAAAAGC AACGCAGAGA TTTCCTAGGT 2880
GAAGCAAGTA TCATGGGACA GTTTGATCAT CCTAACATCA TCCATTTAGA AGGTGTGGTG 2940
ACCAAAAGTA AACCAGTGAT GATCGTGACA GAGTATATGG AGAATGGCTC TTTAGATACA 3000
TTTTTGAAGA AAAACGATGG GCAGTTCACT GTGATTCAGC TTGTTGGCAT GCTGAGAGGT 3060
ATCTCTGCAG GAATGAAGTA CCTTTCTGAC ATGGGCTATG TGCATAGAGA TCTTGCTGCC 3120
AGAAACATCT TAATCAACAG TAACCTTGTG TGCAAAGTGT CTGACTTTGG ACTTTCCCGG 3180
GTACTGGAAG ATGATCCCGA GGCAGCCTAC ACCACAAGGG GAGGAAAAAT TCCAATCAGA 3240
TGGACTGCCC CAGAAGCAAT AGCTTTCCGA AAGTTTACTT CTGCCAGTGA TGTCTGGAGT 3300
TATGGAAATG TAATGTGGGA AGTTGTGTCT TATGGAGAGA GACCCTACTG GGAGGATGACC 3360
AATCAAGATG TGATTAAAGC GGTAGAGGAA GGCTATCGTC TGCCAAGCCC CATGGATTGT 3420
CCTGCTGCTC TCTATCAGTT AATGCTGGAT TGCTGGCAGA AAGAGCGAAA TAGCAGGCCC 3480
AAGTTTGATG AAATAGTCAA CATGTTGGAC AAGCTGATAC GTAACCCAAG TAGTCTGAAG 3540
ACGCTGGTTA ATGCATCCTG CAGAGTATCT AATTTATTGG CAGAACATAG CCCACTAGGA 3600
TCTGGGGCCT ACAGATCAGT AGGTGAATGG CTAGAGGCAA TCAAGATGGG CCGGTATACA 3660
GAGATTTTCA TGGAAAATGG ATACAGTTCA ATGGACGCTG TGGCTCAGGT GACCTTGGAG 3720
GATTTGAGAC GGCTTGGAGT GACTCTTGTC GGTCACCAGA AGAAGATCAT GAACAGCCTT 3780
CAAGAAATGA AGGTGCAGCT GGTAAACGGA ATGGTGCCAT TGT_AA_CTTCA TGTAAATGTC 3840
GCTTCTTCAA GTGAATGATT CTGCACTTTG TAAACAGCAC TGAGATTTAT TTTAACAAAA 3900
AAA

SEQ ID NO:138 PFH3 Protein sequence:
Protein Accession #: CAA64700.1

```
1    11   21   31   41   51
|    |    |    |    |    |
MRGSGPRGAG HRRPPSGGGD TPITPASLAG CYSAPRRAPL WTCLLLCAAL RTLLASPSNE  60
VNLLDSRTVM GDLGWIAFPK NGWEEIGEVD ENYAPIHTYQ VCKVMEQNQN NWLLTSWISN 120
EGASRIFIEL KFTLRDCNSL PGGLGTCKET FNMYYFESDD QNGRNIKENQ YIKIDTIAAD 180
ESFTELDLGD RVMKLNTEVR DVGPLSKKGF YLAFQDVGAC IALVSVRVYY KKCPSVVRHL 240
AVFPDTITGA DSSQLLEVSG SCVNHSVTDE PPKMHCSAEG EWLVPIGKCM CKAGYEEKNG 300
TCQVCRPGFF KASPHIQSCG KCPPHSYTHE EASTSCVCEK DYFRRESDPP TMACTRPPSA 360
PRNAISNVNE TSVFLEWIPP ADTGGRKDVS YYIACKKCNS HAGVCEECGG HVRYLPRQSG 420
LKNTSVMMVD LLAHTNYTFE IEAVNGVSDL SPGARQYVSV NVTTNQAAPS PVTNVKKGKI 480
AKNSISLSWQ EPDRPNGIIL EYEIKHFEKD QETSYTIIKS KETTITAEGL KPASVYVFQI 540
RARTAAGYGV FSRRFEFETT PVFAASSDQS QIPVIAVSVT VGVILLAVVI GVLLSGSCCE 600
CGCGRASSLC AVAHPILIWR CGYSKAKQDP EEEKMHFHNG HIKLPGVRTY IDPHTYEDPN 660
QAVHEFAKEI EASCITIERV IGAGEFGEVC SGRLKLPGKR ELPVAIKTLK VGYTEKQRRD 720
FLGEASIMGQ FDHPNIIHLE GVVTKSKPVM IVTEYMENGS LDTFLKKNDG QFTVIQLVGM 780
LRGISAGMKY LSDMGYVHRD LAARNILINS NLVCKVSDFG LSRVLEDDPE AAYTTRGGKI 840
PIRWTAPEAI AFRKFTSASD VWSYGIVMWE VVSYGERPYW EMTNQDVIKA VEEGYRLPSP 900
MDCPAALYQL MLDCWQKERN SRPKFDEIVN MLDKLIRNPS SLKTLVNASC RVSNLLAEHS 960
PLGSGAYRSV GEWLEAIKMG RYTEIFMENG YSSMDAVAQV TLEDLRRLGV TLVGHQKKIM 1020
NSLQEMKVQL VNGMVPL
```

SEQ ID NO:139 PFH2 DNA SEQUENCE

Nucleic Acid Accession #: NM_016029
Coding sequence: 78-1097 (underlined sequences correspond to start and stop codons)

```
1    11   21   31   41   51
|    |    |    |    |    |
CTGCGATCCC GCAGGGCAGC GACGCGACTC TGGTGCGGGC CGTCTTCTTC CCCCCGAGCT  60
GGGCGTGCGC GGCCGCA_ATG_ AACTGGGAGC TGCTGCTGTG GCTGCTGGTG CTGTGCGCGC 120
TGCTCCTGCT CTTGGTGCAG CTGCTGCGCT TCCTGAGGGC TGACGGCGAC CTGACGCTAC 180
TATGGGCCGA GTGGCAGGGA CGACGCCCAG AATGGGAGCT GACTGATATG GTGGTGTGGG 240
TGACTGGAGC CTCGAGTGGA ATTGGTGAGG AGCTGGCTTA CCAGTTGTCT AAACTAGGAG 300
TTTCTCTTGT GCTGTCAGCC AGAAGAGTGC ATGAGCTGGA AAGGGTGAAA AGAAGATGCC 360
TAGAGAATGG CAATTTAAAA GAAAAAGATA TACTTGTTTT GCCCCTTGAC CTGACCGACA 420
CTGGTTCCCA TGAAGCGGCT ACCAAAGCTG TTCTCCAGGA GTTTGGTAGA ATCGACATTC 480
TGGTCAACAA TGGTGGAATG TCCCAGCGTT CTCTGTGCAT GGATACCAGC TTGGATGTCT 540
ACAGAAAGCT AATAGAGCTT AACTACTTAG GGACGGTGTC CTTGACAAAA TGTGTTCTGC 600
CTCACATGAT CGAGAGGAAG CAAGGAAAGA TTGTTACTGT GAATAGCATC CTGGGTATCA 660
TATCTGTACC TCTTTCCATT GGATACTGTG CTAGCAAGCA TGCTCTCCGG GGTTTTTTTA 720
ATGGCCTTCG AACAGAACTT GCCACATACC CAGGTATAAT AGTTTCTAAC ATTTGCCCAG 780
GACCTGTGCA ATCAAATATT GTGGAGAATT CCCTAGCTGG AGAAGTCACA AAGACTATAG 840
GCAATAATGG AGACCAGTCC CACAAGATGA CAACCAGTCG TTGTGTGCGG CTGATGTTAA 900
TCAGCATGGC CAATGATTTG AAAGAAGTTT GGATCTCAGA ACAACCTTTC TTGTTAGTAA 960
CATATTTGTG GCAATACATG CCAACCTGGG CCTGGTGGAT AACCAACAAG ATGGGGAAGA 1020
AAAGGATTGA GAACTTTAAG AGTGGTGTGG ATGCAGACTC TTCTTATTTT AAAATCTTTA 1080
AGACAAAACA TGAC_TGA_AAA GAGCACCTGT ACTTTTCAAG CCACTGGAGG GAGAAATGGA 1140
AAACATGAAA ACAGCAATCT TCTTATGCTT CTGAATAATC AAAGACTAAT TTGTGATTTT 1200
```

ACTTTTTAAT AGATATGACT TTGCTTCCAA CATGGAATGA AATAAAAAAT AAATAATAAA 1260
AGATTGCCAT GAATCTTGCA AA

SEQ ID NO:140 PFH2 Protein sequence:
Protein Accession #:   NP_057113.1

```
1    11   21   31   41   51
|    |    |    |    |    |
MNWELLLWLL VLCALLLLLV QLLRFLRADG DLTLLWAEWQ GRRPEWELTD MVVWVTGASS  60
GIGEELAYQL SKLGVSLVLS ARRVHELERV KRRCLENGNL KEKDILVLPL DLTDTGSHEA  120
ATKAVLQEFG RIDILVNNGG MSQRSLCMDT SLDVYRKLIE LNYLGTVSLT KCVLPHMIER  180
KQGKIVTVNS ILGIISVPLS IGYCASKHAL RGFFNGLRTE LATYPGIIVS NICPGPVQSN  240
IVENSLAGEV TKTIGNNGDQ SHKMTTSRCV RLMLISMAND LKEVWISEQP FLLVTYLWQY  300
MPTWAWWITN KMGKKRIENF KSGVDADSSY FKIFKTKHD
```

SEQ ID NO:141 PFH1 DNA SEQUENCE
Nucleic Acid Accession #:   NM_021614
Coding sequence: 1-1740 (underlined sequences correspond to start and stop codons)

```
1    11   21   31   41   51
|    |    |    |    |    |
ATGAGCAGCT GCAGGTACAA CGGGGGCGTC ATGCGGCCGC TCAGCAACTT GAGCGCGTCC  60
CGCCGGAACC TGCACGAGAT GGACTCAGAG GCGCAGCCCC TGCAGCCCCC CGCGTCTGTC  120
GGAGGAGGTG GCGGCGCGTC CTCCCCGTCT GCAGCCGCTG CCGCCGCCGC CGCTGTTTCG  180
TCCTCAGCCC CCGAGATCGT GGTGTCTAAG CCCGAGCACA ACAACTCCAA CAACCTGGCG  240
CTCTATGGAA CCGGCGGCGG AGGCAGCACT GGAGGAGGCG GCGGCGGTGG CGGGAGCGGG  300
CACGGCAGCA GCAGTGGCAC CAAGTCCAGC AAAAAGAAAA ACCAGAACAT CGGCTACAAG  360
CTGGGCCACC GGCGCGCCCT GTTCGAAAAG CGCAAGCGGC TCAGCGACTA CGCGCTCATC  420
TTCGGCATGT TCGGCATCGT GGTCATGGTC ATCGAGACCG AGCTGTCGTG GGGCGCCTAC  480
GACAAGGCGT CGCTGTATTC CTTAGCTCTG AAATGCCTTA TCAGTCTCTC CACGATCATC  540
CTGCTCGGTC TGATCATCGT GTACCACGCC AGGGAAATAC AGTTGTTCAT GGTGGACAAT  600
GGAGCAGATG ACTGGAGAAT AGCCATGACT TATGAGCGTA TTTTCTTCAT CTGCTTGGAA  660
ATACTGGTGT GTGCTATTCA TCCCATACCT GGGAATTATA CATTCACATG GACGGCCCGG  720
CTTGCCTTCT CCTATGCCCC ATCCACAACC ACCGCTGATG TGGATATTAT TTTATCTATA  780
CCAATGTTCT TAAGACTCTA TCTGATTGCC AGAGTCATGC TTTTACATAG CAAACTTTTC  840
ACTGATGCCT CCTCTAGAAG CATTGGAGCA CTTAATAAGA TAAACTTCAA TACACGTTTT  900
GTTATGAAGA CTTTAATGAC TATATGCCCA GGAACTGTAC TCTTGGTTTT TAGTATCTCA  960
TTATGGATAA TTGCCGCATG GACTGTCCGA GCTTGTGAAA GGTACCATGA TCAACAGGAT 1020
GTTACTAGCA ACTTCCTTGG AGCGATGTGG TTGATATCAA TAACTTTTCT CTCCATTGGT 1080
TATGGTGACA TGGTACCTAA CACATACTGT GGAAAAGGAG TCTGCTTACT TACTGGAATT 1140
ATGGGTGCTG GTTGCACAGC CCTGGTGGTA GCTGTAGTGG CAAGGAAGCT AGAACTTACC 1200
AAAGCAGAAA AACACGTGCA CAATTTCATG ATGGATACTC AGCTGACTAA AAGAGTAAAA 1260
AATGCAGCTG CCAATGTACT CAGGGAAACA TGGCTAATTT ACAAAAATAC AAAGCTAGTG 1320
AAAAAGATAG ATCATGCAAA AGTAAGAAAA CATCAACGAA AATTCCTGCA AGCTATTCAT 1380
CAATTAAGAA GTGTAAAAAT GGAGCAGAGG AAACTGAATG ACCAAGCAAA CACTTTGGTG 1440
GACTTGGCAA AGACCCAGAA CATCATGTAT GATATGATTT CTGACTTAAA CGAAAGGAAT 1500
GAAGACTTCG AGAAGAGGAT TGTTACCCTG GAAACAAAAC TAGAGACTTT GATTGGTAGC 1560
ATCCACGCCC TCCCTGGGCT CATAAGCCAG ACCATCAGGC AGCAGCAGAG AGATTTCATT 1620
GAGGCTCAGA TGGAGAGCTA CGACAAGCAC GTCACTTACA ATGCTGAGCG GTCCCGGTCC 1680
TCGTCCAGGA GGCGGCGGTC CTCTTCCACA GCACCACCAA CTTCATCAGA GAGTAGCTAG
```

SEQ ID NO:142 PFH1 Protein sequence:
Protein Accession #:   NP_067627

```
1    11   21   31   41   51
|    |    |    |    |    |
MSSCRYNGGV MRPLSNLSAS RRNLHEMDSE AQPLQPPASV GGGGGASSPS AAAAAAAAVS  60
SSAPEIVVSK PEHNNSNNLA LYGTGGGGST GGGGGGGGSG HGSSSGTKSS KKKNQNIGYK  120
LGHRRALFEK RKRLSDYALI FGMFGIVVMV IETELSWGAY DKASLYSLAL KCLISLSTII  180
LLGLIIVYHA REIQLFMVDN GADDWRIAMT YERIFFICLE ILVCAIHPIP GNYTFTWTAR  240
LAFSYAPSTT TADVDIILSI PMFLRLYLIA RVMLLHSKLF TDASSRSIGA LNKINFNTRF  300
VMKTLMTICP GTVLLVFSIS LWIIAAWTVR ACERYHDQQD VTSNFLGAMW LISITFLSIG  360
YGDMVPNTYC GKGVCLLTGI MGAGCTALVV AVVARKLELT KAEKHVHNFM MDTQLTKRVK  420
NAAANVLRET WLIYKNTKLV KKIDHAKVRK HQRKFLQAIH QLRSVKMEQR KLNDQANTLV  480
DLAKTQNIMY DMISDLNERS EDFEKRIVTL ETKLETLIGS IHALPGLISQ TIRQQQRDFI  540
EAQMESYDKH VTYNAERSRS SSRRRRSSST APPTSSESS
```

SEQ ID NO:143 PFG9 DNA SEQUENCE
Nucleic Acid Accession #:   AL110139, coding region is FGENESH predicted
Coding sequence: 1-1896 (underlined sequences correspond to start and stop codons)

```
1    11   21   31   41   51
|    |    |    |    |    |
```

355

```
 |    |    |    |    |    |
ATGCGCGCCG TGCCGCTGCC CGCCCCGCTC CTGCCGCTGC TGCTGCTCGC GCTCCTGGCC   60
GCTCCCGCCG CCCGCGCCAG CAGAGCCGAG TCCGTCTCCG CGCCGTGGCC CGAACCCGAG  120
CGCGAGTCGC GGCCACCGCC CGGCCCGGGG CCCGGGAACA CCACCCGGTT TGGGTCTGGG  180
GCGGCGGGCG GCAGCGGCAG CTCCAGCTCC AACAGCAGTG GCGACGCCTT GGTGACCCGC  240
ATTTCCATCC TCCTCCGCGA CCTACCCACC CTCAAGGCAG CCGTGATCGT GGCGTTCGCC  300
TTTACCACCC TCCTCATCGC CTGCCTGCTG CTGCGCGTCT TCAGGTCGGG AAAGAGGTTA  360
AAGAAGACAC GCAAGTATGA TATCATCACC ACTCCAGCAG GCGAGTGGA AATGGCGCCA   420
CTAAATGAAG AGGATGATGA AGATGAGGAC TCCACAGTAT TCGACATCAA ATACAGAGTG  480
TCCTTGCCGG CTGCACTGAG ACGTCAGCTG CCAGGGTGCC AGACGCTACT GACAGTTCCT  540
GTGCCCCCAC CCTTCATCCT CGACATTGAC CTTCCAGCAA GATGCAGTGG AAGGCCTGAT  600
GGTGGAATCA GACCTGGTAA AACCTGTTTC CCAGCCTGGT GGCATCCTGT GGAAAGTTGG  660
TCAGCTGCAA CCTGGGGTGT GAAGGACTGG ACCTGGAAGC CCTCTTGCGT CGGAGGTGTT  720
GAAACCAAAA CGAACGTTAT GTATAAAACC CCAGCTCCAT CGTGCGTGTC AGGCATCTGC  780
TCAGACTGTC ACTGGCAAGC TCGTTTCCAC GTCACCACAA TGGAGTTGCT TCTGCCACCC  840
TTTGGGCATC CCTTTAAAGT GCCCCCTACT TCTACTCCCC ATGGTTTTCG ACAACTGCAG  900
CTGAATCTCA TGGAAAAGCT GGATTCCTCT GCCTTACGCA GAAACACCCG GGCTCCATCT  960
GCCAGGTGCT TGCCACTGGT CCTGGCAGAA ATGGCGGCTG CTGAAAGTGA CCTTCCAAAT 1020
CCTTGGTGGC ACTTCAGCGC CACAGGCTCT CCAATAAAAA CCCTTTACAC ACAAACCATG 1080
AGTACCTTGG GCTTGGATGT TTTCTGTGGT GCCGGCCAGC GGGGCACCTT TTGTGAAGAC 1140
AGAGCAGTGA CTAAGGTTCT CCAGGGTAGC TCTTTCTCCA AACAGCTGCG CTGGAAGCCA 1200
GCCCTAGAGA GTGGGTTTCC CCATCATCTC AGGCTTCTCA GGGAGTGTCC TCCGCTGAGC 1260
ACCCATCCTG TCAGGTTGGC TCGTTCAGAT GCCCGGGGAC AAGCCAGCCT GACGGGGAGG 1320
AGGGTGTTTC GGCGTCCGCG GCAGTCTCTG CATGGCGGAG GGTCAGCGGG TACCGCAACT 1380
TGCCTTTTGG TTTTGAAGAT TCTGTTGAGG CGCCATCCTC ACCTTGACCT CTTCTACAAA 1440
ATCTGTCTCC CCTGCTGTGC CGTGGAACAC CTACGGGAAG CCAAGAGAAG CTCAGTGACT 1500
GTCCTTGCGT CATTTGAGCA GAGCCCACAA AAGGCAGCTG CTGCCCACGG GGAGCCTGTC 1560
AAACGAGGGC CCAGTGGGCA ATTGACCAGA CACACATGCC CTGGCTGGGG GATCACACAT 1620
GCGAACCTGC AGACAATTCC AGATACCCAA GGCCAGGAGG GCCCACGTGA GGATGTCACT 1680
CACCCTGGAG GAGACTTGGA TGGGGTGGCA AATTTCTATT TGGAGGAAGA GGGTTTCCAG 1740
GATGGCAGAT GCCAGAAGAT GGTCCTGATG TCTGAGGAAG GGCCACCTAG TTTGACAGGA 1800
TGTGAGAGGC TCACAGGTTC CCATCACTTC TCCAGCCATT CCAAGTCTTG GTCCTTCCTT 1860
TCCCCCCGAC AGCCCCTGTT TCTGTCCAGG CCCTGA
```

SEQ ID NO:144 PFG9 Protein sequence:
Protein Accession #:   none available, FGENESH predicted

```
 1    11    21    31    41    51
 |    |    |    |    |    |
MRAVPLPAPL LPLLLLALLA APAARASRAE SVSAPWPEPE RESRPPPGPG PGNTTRFGSG   60
AAGGSGSSSS NSSGDALVTR ISILLRDLPT LKAAVIVAFA FTTLLIACLL LRVFRSGKRL  120
KKTRKYDIIT TPAERVEMAP LNEEDDEDED STVFDIKYRV SLPAALRRQL PGCQTLLTVP  180
VPPPFILDID LPARCSGRPD GGIRPGKTCF PAWWHPVESW SAATWGVKDW TWKPSCVGGV  240
ETKTNVMYKT PAPSCVSGIC SDCHWQARFH VTTMELLLPP FGHPFKVPPT STPHGFRQLQ  300
LNLMEKLDSS ALRRNTRAPS ARCLPLVLAE MAAAESDLPN PWWHFSATGS PIKTLYTQTM  360
STLGLDVFCG AGQRGTFCED RAVTKVLQGS SFSKQLRWKP ALESGFPHHL RLLRECPPLS  420
THPVRLARSD ARGQASLTGR RVFRRPRQSL HGGGSAGTAT CLLVLKILLR RHPHLDLFYK  480
ICLPCCAVEH LREAKRSSVT VLASFEQSPQ KAAAAHGEPV KRGPSGQLTR HTCPGWGITH  540
ANLQTIPDTQ GQEGPREDVT HPGGDLDGVA NFYLEEEGFQ DGRCQKMVLM SEEGPPSLTG  600
CERLTGSHHF SSHSKSWSFL SPRQPLFLSR P
```

SEQ ID NO:145 PFG6 DNA SEQUENCE
Nucleic Acid Accession #: NM_013427
Coding sequence: 875-3799 (underlined sequences correspond to start and stop codons)

```
 1    11    21    31    41    51
 |    |    |    |    |    |
GGCTGGGCTG CGAATAGCGT GTTCCTCTCC GGCGGAACAC ACACACCCGG CCTTGGGGCT   60
GTCTCCTGAA GCTCCCTCCT CCACGGAGAG CGCTGAGCGC CGCCGGGAAT TCCATCCCAC  120
CGTGGGCACG CAGTCTTTGG AGGTCCCGGG CGCAGCACGC TCGGTGTCCC CACACTGCAG  180
CAAGACAGAG ACCCCGCGGG AACCTTGAGC TTGGAACAAC CCTTGAGCCT CTGCAGTCGG  240
AAGAGTGGGC GCAGCAGCCC AGCGGAGGCC AGGCGCGCAA CCTCGGGCGC CGGGGCAAGG  300
AGAGAGTGCA GGGAGGCGCA GCTCAGGCGC CCGGCTCAGG AGCGGGAGGA AGTTCTCGCG  360
GCGCCGGGAG CGCGGTGGAC GCGCCCTGGG CGCACGCCCA GGCAGCCTTC TCCCTGGCCC  420
TCGGGACTGT CCTCGGGCGG CAAGGAGGAG CTTGCTGGAG TCTTAGAGGC CATCCAGAGC  480
CAGCGAGCAG GAGCGCTGCG TCTCCCGCCT CAGCTAGGAA GGGGGAGTGG CGCTGGCAGG  540
CTGGAGCTGG GAACCCAGCG AGCGCCTGAC CTTCCTCCTC CTCTTCCTGA CCCTCTTCGC  600
GTCTTGGGCT CCGGAGGAAG GTTCTAGCGG CTGCAGGAGG TCCCCAGACC CATTTTCCTA  660
GAAGGCTGGT GATGGATCTG CTGCTCCTGC CGCCGCCGGG GCACTGGGAG CGCACCGGCG  720
GCGCGTGAGC TGGGCTTTGC TCTCCACCGC CCTGGGCAAA CCCCGGGCCA GCCCCGCCTG  780
GCACCTTTGC CTGAGTCCCT TTCGGTTCCC GACCCAAAGC CACCAGCGTC CAGGGAGGGA  840
GGAGGAGGTG GTCCTCAGGT GCAGCCCCGC CGAGATGTCC GCGCAGAGCC TGCTCCACAG  900
CGTCTTCTCC TGTTCCTCGC CCGCTTCAAG TAGCGCGGCC TCGGCCAAGG GCTTCTCCAA  960
GAGGAAGCTG CGCCAGACCC GCAGCCTGGA CCCGGCCCTG ATCGGCGGCT GCGGGAGCGA 1020
CGAGGCGGGC GCGGAGGGCA GTGCGCGGGG AGCCACGGGC GGCCGCCTCT ACTCCCCATC 1080
ACTCCCAGCC GAGAGTCTCG GCCCTCGCTT GGCGTCCTCT TCCCGGGGTC CGCCCCCCAG 1140
GGCCACCAGG CTACCGCCTC CTGGACCTCT TTGCTCGTCC TTCTCCACAC CCAGCACCCC 1200
```

```
GCAGGAGAAG TCACCATCCG GCAGCTTTCA CTTTGACTAT GAGGTTCCCC TGGGTCGCGG 1260
CGGCCTCAAG AAGAGCATGG CCTGGGACCT GCCTTCTGTC CTGGCCGGGC CAGCCAGTAG 1320
CCGAAGCGCT TCCAGCATCC TCTGTTCATC CGGGGGGAGGC CCCAATGGCA TCTTCGCTTC 1380
TCCTAGGAGG TGGCTCCAGC AGAGGAAGTT CCAGTCCCCA CCCGACAGTC GCGGGCACCC 1440
CTACGTCGTG TGGAAATCCG AGGGTGATTT CACCTGGAAC AGCATGTCAG GCCGCAGTGT 1500
GCGGCTGAGG TCAGTCCCCA TCCAGAGTCT CTCAGAGCTG GAGAGGGCCC GGCTGCAGGA 1560
AGTGCCTTTT TATCAGTTGC AACAGGACTG TGACCTGAGC TGTCAGATCA CCATTCCCAA 1620
AGATGGACAA AAGAGAAAGA AATCTTTAAG AAAGAAACTG GATTCACTAG GAAAGGAGAA 1680
AAACAAAGAC AAAGAATTCA TCCCACAGGC ATTTGGAATG CCCTTATCCC AAGTCATTGC 1740
GAATGACAGG GCCTATAAAC TCAAGCAGGA CTTGCAGAGG GACGAGCAGA AAGATGCATC 1800
TGACTTTGTG GCTTCCCTCC TCCCATTTGG AAATAAAAGA CAAAACAAAG AACTCTCAAG 1860
CAGTAACTCA TCTCTCAGCT CAACCTCAGA AACACCGAAT GAGTCAACGT CCCCAAACAC 1920
CCCGGAACCG GCTCCTCGGG CTAGGAGGAG GGGTGCCATG TCAGTGGATT CTATCACCGA 1980
TCTTGATGAC AATCAGTCTC GACTACTAGA AGCTTTACAA CTTTCCTTGC CTGCTGAGGC 2040
TCAAAGTAAA AAGGAAAAAG CCAGAGATAA GAAACTCAGT CTGAATCCTA TTTACAGACA 2100
GGTCCCTAGG CTGGTGGACA GCTGCTGTCA GCACCTAGAA AAACATGGCC TCCAGACAGT 2160
GGGGATATTC CGAGTTGGAA GCTCAAAAAA GAGAGTGAGA CAATTACGTG AGGAATTTGA 2220
CCGTGGGATT GATGTCTCTC TGGAGGAGGA GCACAGTGTT CATGATGTGG CAGCCTTGCT 2280
GAAAGAGTTC CTGAGGGACA TGCCAGACCC CCTTCTCACC AGGGGAGCTGT ACACAGCTTT 2340
CATCAACACT CTCTTGTTGG AGCCGGAGGA ACAGCTGGGC ACCTTGCAGC TCCTCATATA 2400
CCTTCTACCT CCCTGCAACT GCGACACCCT CCACCGCCTG CTACAGTTCC TCTCCATCGT 2460
GGGCCAGGCT GCCGATGACA ACATCAGCAA AGATGGCCTG GAGGTCACTG GGAATAAAAT 2520
GACATCTCTA AACTTAGCCA CCATATTTGG ACCCAACCTG CTGCACAAGC AGAAGTCATC 2580
AGACAAAGAA TTCTCAGTTC AGAGTTCAGC CCGGGCTGAG GAGAGCACGG CCATCATCGC 2640
TGTTGTGCAA AAGATGATTG AAAATTATGA AGCCCTGTTC ATGGTTCCCC CAGATCTCCA 2700
GAACGAAGTG CTGATCAGCC TGTTAGAGAC CGATCCTGAT GTCGTGGACT ATTTACTCAG 2760
AAGAAAGGCT TCCCAATCAT CAAGCCCTGA CATGCTGCAG TCGGAAGTTT CCTTTTCCGT 2820
GGGAGGGAGG CATTCATCTA CAGACTCCAA CAAGGCCTCC AGCGGAGACA TCTCCCCTTA 2880
TGACAACAAC TCCCCAGTGC TGTCTGAGCG CTCCCTGCTG GCTATGCAAG AGGACGCGGT 2940
CCCCGGGGGGC TCGGAGAAGC TTTACAGAGT GCCAGGGCAG TTTATGCTGG TGGGCCACTT 3000
GTCGTCGTCA AAGTCAAGGG AAAGTTCTCC TGGACCAAGG CTTGGGAAAG ATCTGTCAGA 3060
GGAGCCTTTC GATATCTGGG GAACTTGGCA TTCAACATTA AAAAGCGGAT CCAAAGACCC 3120
AGGAATGACA GGTTCCTCTG GAGACATTTT TGAAAGCAGC TCCCTAAGAG CGGGGCCCTG 3180
CTCCCTTTCT CAAGGGAACC TGTCCCCAAA TTGGCCTCGG TGGCAGGGGA GCCCCGCAGA 3240
GCTGGACAGC GACACGCAGG GGGCTCGGAG GACTCAGGCC GCAGCCCCCG CGACGGAGGG 3300
CAGGGCCCAC CCTGCGGTGT CGCGCGCCTG CAGCACGCCC CACGTCCAGG TGGCAGGGAA 3360
AGCCGAGCGG CCCACGGCCA GGTCGGAGCA GTACTTGACC CTGAGCGGCG CCCACGACCT 3420
CAGCGAGAGT GAGCTGGATG TGGCCGGGCT GCAGAGCCGG GCCACACCTC AGTGCCAAAG 3480
ACCCCATGGG AGTGGGAGGG ATGACAAGCG GCCCCCGCCT CCATACCCGG GCCCAGGGAA 3540
GCCCGCGGCA GCGGCAGCCT GGATCCAGGG GCCCCCGGAA GGCGTGGAGA CACCCACGGA 3600
CCAGGGAGGC CAAGCAGCCG AGCGAGAGCA GCAGGTCACG CAGAAAAAAC TGAGCAGCGC 3660
CAACTCCCTG CCAGCGGGCG AGCAGGACAG TCCGCGCCTG GGGGACGCTG GCTGGCTCGA 3720
CTGGCAGAGA GAGCGCTGGC AGATCTGGGA GCTCCTGTCG ACCGACAACC CCGATGCCCT 3780
GCCCGAGACG CTGGTCTGAG CCCGCACCCA GCCGAGCCCC CCCTGCCCCG AGCCCCCCGC 3840
CCTCCAGCCC AGGGGGGACC GTGGGTGGTG GCCACTGGCA CACTTAGTGT TCTTCTTTCA 3900
CACTTCTCAA AAGTGACACA AGAGAAATCC AGTTCACCTA CAGAGGTAGA GCACTCACCG 3960
CCCCGCCATT GAGAATAAGG TTCCATTGCG TAGCCAGCCT TAGGAAAAAC AAACAGAACC 4020
CAAACCAGAT GGCAATGTCC AATCTAAAAA CGTCCCTCTT GGCTCTATAA TATAAGGATAC 4080
AACTCTTGCT TGGTATAGCC TAACCGTATT TATGTGTCTT CGGTTTTGAC TATTGTGTAT 4140
TCTGTAACAG ATTATGTATA ATCATATATG ATATATTCAC AAAGAGAAAA CAAAAGGAAC 4200
TTTTAAAAAA AAAATCACTT CACTTATATT AAGCAATGAG ATATACTAAA CAATGAGATT 4260
CTATAGAATG TTCTAGAATG TGCACAAGCG GGTTTCTGTG CTTTTGCCAT AGCTTTATAA 4320
CTGGGGATAA CCCTTCCTTC GATACCAAAC ACTAACAAGA GAAGGCAGAA TATGAGAAGC 4380
CATATTTTTA CATAGGAGTC AGATACAAAA AGAAAAATCA CTGAATGCTT TTAGATATTG 4440
AATACGTTTT CAGGAAAATG CTAAATCTGA TAGATTACGA AATATATTTT TAGAACTTGT 4500
TTAGAAAGGA TTCAGTTAAC CAAACAAGAA AAAGGCAGTG CCTCACAAAG AAATTAAGAA 4560
GTTGTCCGTC CCACGTTACA TCAAATTCAG TTTTATATAG GCCATATATA ATATATATTT 4620
ATAATGTATA ATTTTTATGT ATTTTTCAAA ACTACAAACT GGAATCCAAC TATAAAGTGT 4680
TTAAGAATCT ACACAGAATA TTCAAATTAT AGAACATGTT TTTTCCCTTT GCCCCATAAT 4740
CAGTATTTGC CAAATTACAT GCAATTCCTT AAAAACTAAA TCACATTGGT AAAAGGCCTA 4800
CAGCTTTGTA CTTACATTGT GCCAAAGGCT GAGGAAATGT TTTCTTTCGA ATTTTTATGT 4860
GTATTGTAAA ATGTTCTACC GTACTTTAGT AGTTTGAAGT TTTCAAGTGC ATAACTATTT 4920
TTGACCAGCA GAAGGCGATA CGCTTCAGTA TTTTATGCAA TTTTTTTTCA CTTCGAAGGG 4980
AAAGTGTATT ATAAAAAAAG ATTTTTTTTT TTTAAAACAT GCTACTCTTA ATTTTCATGT 5040
TGGTGATGAA ATTCCCAGTG GTGTTTCTTA AGGTTCTATC TTGTGCCATG ATGAATAAAA 5100
AGTTAAGCAA AAAAAAAAAA AAAAAAAAAA AAA
```

SEQ ID NO:146 PFG6 Protein sequence:
Protein Accession #:   NP_038286.1

```
1      11     21     31     41     51
|      |      |      |      |      |
MSAQSLLHSV FSCSSPASSS AASAKGFSKR KLRQTRSLDP ALIGGCGSDE AGAEGSARGA   60
TAGRLYSPSL PAESLGPRLA SSSRGPPPRA TRLPPPGPLC SSFSTPSTPQ EKSPSGSFHF  120
DYEVPLGRGG LKKSMAWDLP SVLAGPASSR SASSILCSSG GGPNGIFASP RRWLQQRKFQ  180
SPPDSRGHPY VVWKSEGDFT WNSMSGRSVR LRSVPIQSLS ELERARLQEV PFYQLQQDCD  240
LSCQITIPKD GQKRKKSLRK KLDSLGKEKN KDKEFIPQAF GMPLSQVIAN DRAYKLKQDL  300
QRDEQKDASD FVASLLPFGN KRQNKELSSS NSSLSSTSET PNESTSPNTP EPAPRARRRG  360
```

AMSVDSITDL DDNQSRLLEA LQLSLPAEAQ SKKEKARDKK LSLNPIYRQV PRLVDSCCQH 420
LEKHGLQTVG IFRVGSSKKR VRQLREEFDR GIDVSLEEEH SVHDVAALLK EFLRDMPDPL 480
LTRELYTAFI NTLLLEPEEQ LGTLQLLIYL LPPCNCDTLH RLLQFLSIVA RHADDNISKD 540
GQEVTGNKMT SLNLATIFGP NLLHKQKSSD KEFSVQSSAR AEESTAIIAV VQKMIENYEA 600
LFMVPPDLQN EVLISLLETD PDVVDYLLRR KASQSSSPDM LQSEVSFSVG GRHSSTDSNK 660
ASSGDISPYD NNSPVLSERS LLAMQEDAAP GGSEKLYRVP GQFMLVGHLS SSKSRESSPG 720
PRLGKDLSEE PFDIWGTWHS TLKSGSKDPG MTGSSGDIFE SSSLRAGPCS LSQGNLSPNW 780
PRWQGSPAEL DSDTQGARRT QAAAPATEGR AHPAVSRACS TPHVQVAGKA ERPTARSEQY 840
LTLSGAHDLS ESELDVAGLQ SRATPQCQRP HGSGRDDKRP PPPYPGPGKP AAAAAWIQGP 900
PEGVETPTDQ GGQAAEREQQ VTQKKLSSAN SLPAGEQDSP RLGDAGWLDW QRERWQIWEL 960
LSTDNPDALP ETLV

SEQ ID NO:147 PFG4 DNA SEQUENCE

Nucleic Acid Accession #: NM_002202
Coding sequence: 240-1289 (underlined sequences correspond to start and stop codons)

```
1    11   21   31   41   51
|    |    |    |    |    |
CCCCCGAGCC GCGCCGAGTC TGCCGCCGCC GCAGCGCCTC CGCTCCGCCA ACTCCGCCGG   60
CTTAAATTGG ACTCCTAGAT CCGCGAGGGC GCGGCGCAGC CGAGCAGCGG CTCTTTCAGC  120
ATTGGCAACC CCAGGGGCCA ATATTTCCCA CTTAGCCACA GCTCCAGCAT CCTCTCTGTG  180
GGCTGTTCAC CAACTGTACA ACCACCATTT CACTGTGGAC ATTACTCCCT CTTACAGATA  240
TGGGAGACAT GGGAGATCCA CCAAAAAAAA AACGTCTGAT TTCCCTATGT GTTGGTTGCG  300
GCAATCAGAT TCACGATCAG TATATTCTGA GGGTTTCTCC GGATTTGGAA TGGCATGCGG  360
CATGTTTGAA ATGTGCGGAG TGTAATCAGT ATTTGGACGA GAGCTGTACA TGCTTTGTTA  420
GGGATGGGAA AACCTACTGT AAAAGAGATT ATATCAGGTT GTACGGGATC AAATGCGCCA  480
AGTGCAGCAT CGGCTTCAGC AAGAACGACT TCGTGATGCG TGCCCGCTCC AAGGTGTATC  540
ACATCGAGTG TTTCCGCTGT GTGGCCTGCA GCCGCCAGCT CATCCCTGGG GACGAATTTG  600
CGCTTCGGGA GGACGGTCTC TTCTGCCGAG CAGACCACGA TGTGGTGGAG AGGGCCAGTC  660
TAGGCGCTGG CGACCCGCTC AGTCCCCTGC ATCCAGCGCG GCCACTGCAA ATGGCAGCGG  720
AGCCCATCTC CGCCAGGCAG CCAGCCCTGC GGCCCCACGT CCACAAGCAG CCGGAGAAGA  780
CCACCCGCGT GCGGACTGTG CTGAACGAGA AGCAGCTGCA CACCTTGCGG ACCTGCTACG  840
CCGCAAACCC GCGGCCAGAT GCGCTCATGA AGGAGCAACT GGTAGAGATG ACGGGCCTCA  900
GTCCCCGTGT GATCCGGGTC TGGTTTCAAA ACAAGCGGTG CAAGGACAAG AAGCGAAGCA  960
TCATGATGAA GCAACTCCAG CAGCAGCAGC CCAATGACAA AACTAATATC CAGGGGATGA 1020
CAGGAACTCC CATGGTGGCT GCCAGTCCAG AGAGACACGA CGGTGGCTTA CAGGCTAACC 1080
CAGTGGAAGT ACAAAGTTAC CAGCCACCTT GGAAAGTACT GAGCGACTTC GCCTTGCAGA 1140
GTGACATAGA TCAGCCTGCT TTTCAGCAAC TGGTCAATTT TTCAGAAGGA GGACCGGGCT 1200
CTAATTCCAC TGGCAGTGAA GTAGCATCAA TGTCCTCTCA ACTTCCAGAT ACACCTAACA 1260
GCATGGTAGC CAGTCCTATT GAGGCATGAG GAACATTCAT TCTGTATTTT TTTTCCCTGT 1320
TGGAGAAAGT GGGAAATTAT AATGTCGAAC TCTGAAACAA AAGTATTTAA CGACCCAGTC 1380
AATGAAAACT GAATCAAGAA ATGAATGCTC CATGAAATGC ACGAAGTCTG TTTTAATGAC 1440
AAGGTGATAT GGTAGCAACA CTGTGAAGAC AATCATGGGA TTTTACTAGA ATTAAACAAC 1500
AAACAAAACG CAAAACCCAG TATATGCTAT TCAATGATCT TAGAAGTACT GAAAAAAAAA 1560
GACGTTTTTA AAACGTAGAG GATTTATATT CAAGGATCTC AAAGAAAGCA TTTTCATTTC 1620
ACTGCACATC TAGAGAAAAA CAAAAATAGA AAATTTTCTA GTCCATCCTA ATCTGAATGG 1680
TGCTGTTTCT ATATTGGTCA TTGCCTTGCC AAACAGGAGC TCCAGCAAAA GCGCAGGAAG 1740
AGAGACTGGC CTCCTTGGCT GAAAGAGTCC TTTCAGGAAG GTGGAGCTGC ATTGGTTTGA 1800
TATGTTTAAA GTTGACTTTA ACAAGGGGTT AATTGAAATC CTGGGTCTCT TGGCCTGTCC 1860
TGTAGCTGGT TTATTTTTTA CTTTGCCCCC TCCCCACTTT TTTTGAGATC CATCCTTTAT 1920
CAAGAAGTCT GAAGCGACTA TAAAGGTTTT TGAATTCAGA TTTAAAAACC AACTTATAAA 1980
GCATTGCAAC AAGGTTACCT CTATTTTGCC ACAAGCGTCT CGGGATTGTG TTTGACTTGT 2040
GTCTGTCCAA GAACTTTTCC CCCAAAGATG TGTATAGTTA TTGGTTAAAA TGACTGTTTT 2100
CTCTCTCTAT GGAAATAAAA AGGAAAAAAA AAAGGAAACT TTTTTTGTTT GCTCTTGCAT 2160
TGCAAAAATT ATAAAGTAAT TTATTATTTA TTGTCGGAAG ACTTGCCACT TTTCATGTCA 2220
TTTGACATTT TTTGTTTGCT GAAGTGAAAA AAAAAGATAA AGGTTGTACG GTGGTCTTTG 2280
AATTATATGT CTAATTCTAT GTGTTTTGTC TTTTTCTTAA ATATTATGTG AAATCAAAGC 2340
GCCCATATGTA GAATTATATC TTCAGGACTA TTTCACTAAT AAACATTTGG CATAGAT
```

SEQ ID NO:148 PFG4 Protein sequence:
Protein Accession #: NP_002193.1

```
1    11   21   31   41   51
|    |    |    |    |    |
MGDPPKKKRL ISLCVGCGNQ IHDQYILRVS PDLEWHAACL KCAECNQYLD ESCTCFVRDG   60
KTYCKRDYIR LYGIKCAKCS IGFSKNDFVM RARSKVYHIE CFRCVACSRQ LIPGDEFALR  120
EDGLFCRADH DVVERASLGA GDPLSPLHPA RPLQMAAEPI SARQPALRPH VHKQPEKTTR  180
VRTVLNEKQL HTLRTCYAAN PRPDALMKEQ LVEMTGLSPR VIRVWFQNKR CKDKKRSIMM  240
KQLQQQQPND KTNIQGMTGT PMVAASPERH DGGLQANPVE VQSYQPPWKV LSDFALQSDI  300
DQPAFQQLVN FSEGGPGSNS TGSEVASMSS QLPDTPNSMV ASPIEA
```

358

SEQ ID NO:149 PFG2 DNA SEQUENCE

Nucleic Acid Accession #: NM_001172
Coding sequence: 39-1103 (underlined sequences correspond to start and stop codons)

```
 1    11    21    31    41    51
 |    |     |     |     |     |
GCGGAGCTCT GCCTTGGAGA TTCTCAGTGC TGCGGATCAT GTCCCTAAGG GGCAGCCTCT   60
CGCGTCTCCT CCAGACGCGA GTGCATTCCA TCCTGAAGAA ATCCGTCCAC TCCGTGGCTG  120
TGATAGGAGC CCCGTTCTCA CAAGGGCAGA AAAGAAAAGG AGTGGAGCAT GGTCCCGCTG  180
CCATAAGAGA AGCTGGCTTG ATGAAAAGGC TCTCCAGTTT GGGCTGCCAC CTAAAAGACT  240
TTGGAGATTT GAGTTTTACT CCAGTCCCCA AAGATGATCT CTACAACAAC CTGATAGTGA  300
ATCCACGCTC AGTGGGTCTT GCCAACCAGG AACTGGCTGA GGTGGTTAGC AGAGCTGTGT  360
CAGATGGCTA CAGCTGTGTC ACACTGGGAG GAGACCACAG CCTGGCAATC GGTACCATTA  420
GTGGCCATGC CCGACACTGC CCAGACCTTT GTGTTGTCTG GGTTGATGCC CATGCTGACA  480
TCAACACACC CCTTACCACT TCATCAGGAA ATCTCCATGG ACAGCCAGTT TCATTTCTCC  540
TCAGAGAACT ACAGGATAAG GTACCACAAC TCCCAGGATT TTCCTGGATC AAACCTTGTA  600
TCTCTTCTGC AAGTATTGTG TATATTGGTC TGAGAGACGT GGACCCTCCT GAACATTTTA  660
TTTTAAAGAA CTATGATATC CAGTATTTTT CCATGAGAGA TATTGATCGA CTTGGTATCC  720
AGAAGGTCAT GGAACGAACA TTTGATCTGC TGATTGGCAA GAGACAAAGA CCAATCCATT  780
TGAGTTTTGA TATTGATGCA TTTGACCCTA CACTGGCTCC AGCCACAGGA ACTCCTGTTG  840
TCGGGGGACT AACCTATCGA GAAGGCATGT ATATTGCTGA GGAAATACAC AATACAGGGT  900
TGCTATCAGC ACTGGATCTT GTTGAAGTCA ATCCTCAGTT GGCCACCTCA GAGGAAGAGG  960
CGAAGACTAC AGCTAACCTG GCAGTAGATG TGATTGCTTC AAGCTTTGGT CAGACAAGAG 1020
AAGGAGGGCA TATTGTCTAT GACCAACTTC CTACTCCCAG TTCACCAGAT GAATCAGAAA 1080
ATCAAGCACG TGTGAGAATT TAGGAGACAC TGTGCACTGA CATGTTTCAC AACAGGCATT 1140
CCAGAATTAT GAGGCATTGA GGGGATAGAT GAATACTAAA TGGTTGTCTG GGTCAATACT 1200
GCCTTAATGA GAACATTTAC ACATTCTCAC AATTGTAAAG TTTCCCCTCT ATTTTGGTGA 1260
CCAATACTAC TGTAAATGTA TTTGGTTTTT TGCAGTTCAC AGGGTATTAA TATGCTACAG 1320
TACTATGTAA ATTTAAAGAA GTCATAAACA GCATTTATTA CCTTGGTATA TCATACTGGT 1380
CTTGTTGCTG TTGTTCCTTC ACATTTAAGT GGTTTTTCAT CTTTCCTCCC TCCTCCCACA 1440
GCCTGGCTAT ACAGTGCATC CTTGAACTGT CAGCCCACAG CAGCAATATG CTTATTCTAT 1500
CCACATCCCT AACATCATGC ATTCACAAGG TCAAAGTTCT GGTCCACAAA CCCTTCCCTA 1560
TAGAAGTTCA ATGGCTGCGA AAGAATTTGT AGTAAACCAG GCCTCCCAGG ATGGCGAGCT 1620
CCAGTAAGAT GATAATGGAA AGCAGCAGCT TGTTGGTTGT CACTCTACAA AGAGAAGCAA 1680
AGTGGGGAGT AGTCAGAAGT TTGGATAACC TTCCTTCTAA ACATTTGGGG GTTAGACCTG 1740
GGACCACGGC TGGATACTCT GAGGCTGTAT GTTTGATCAC ACAGCCACTT AGCAGGAAGT 1800
ACTCATAAGG TTCTTTAGCT GTCACTTAGG GATAAACACTG TCTACCTCAC AGAAATGTTA 1860
AACTGAGACA ATAAAACCCA AAGCAT
```

SEQ ID NO:150 PFG2 Protein sequence:
Protein Accession #: NP_001163.1

```
 1    11    21    31    41    51
 |    |     |     |     |     |
MSLRGSLSRL LQTRVHSILK KSVHSVAVIG APFSQGQKRK GVEHGPAAIR EAGLMKRLSS   60
LGCHLKDFGD LSFTPVPKDD LYNNLIVNPR SVGLANQELA EVVSRAVSDG YSCVTLGGDH  120
SLAIGTISGH ARHCPDLCVV WVDAHADINT PLTTSSGNLH GQPVSFLLRE LQDKVPQLPG  180
FSWIKPCISS ASIVYIGLRD VDPPEHFILK NYDIQYFSMR DIDRLGIQKV MERTFDLLIG  240
KRQRPIHLSF DIDAFDPTLA PATGTPVVGG LTYREGMYIA EEIHNTGLLS ALDLVEVNPQ  300
LATSEEEAKT TANLAVDVIA SSFGQTREGG HIVYDQLPTP SSPDESENQA RVRI
```

SEQ ID NO:151 PFG1 DNA SEQUENCE

Nucleic Acid Accession #: NM_017906
Coding sequence: 80-1255 (underlined sequences correspond to start and stop codons)

```
 1    11    21    31    41    51
 |    |     |     |     |     |
AATTATATAT TTTTACTCTA TGTTTCTCTA CATGTTTTTT TCTTTCCGTT GCTGGCGGAA   60
GAGGCACGTG CGCTGCTGAA TGGAGCTGGT CGCTGGTTGC TACGAGCAGG TCCTCTTTGG  120
GTTCGCTGTA CACCCGGAGC CCAAGGCTTG CGGCGACCAC GAGCAATGGA CTCTTGTGGC  180
TGACTTCACT CACCATGCTC ACACTGCCTC CTTGTCAGCA GTAGCTGTAA ATAGTCGTTT  240
TGTGGTCACT GGGAGCAAAG ATGAAACAAT TCACATTTAT GACATGAAAA AGAAGATTGA  300
GCATGGGGCT CTAGTGCATC ACAGTGGTAC AATAACTTGC CTGAAATTCT ATGGCAACAG  360
GCATTTAATC AGTGGAGCGG AAGATGGACT CATCTGTATC TGGGATGCAA AGAAATGGGA  420
ATGCCTGAAG TCAATTAAAG CTCACAAAGG ACAGGTGACC TTCCTTTCTA TTCACCCATC  480
TGGCAAGTTG GCCCTGTCGG TTGGTACAGA TAAAACTTTA AGAACGTGGA ATCTTGTAGA  540
AGGAAGATCA GCATTCATAA AAAATATAAA ACAAAATGCT CACATAGTAG AATGGTCCCC  600
AAGAGGAGAG CAGTATGTAG TTATCATACA GAATAAAATA GACATCTATC AGCTTGACAC  660
TGCATCCATT AGTGGCACCA TCACAAATGA AAAGAGAATT CCTCTGTTA AATTTCTTTC  720
AGAGTCTGTC CTTGCAGTGG CTGGAGATGA AGAAGTTATA AGGTTTTTTG ACTGTGATTC  780
ACTAGTGTGC CTCTGCGAAT TTAAAGCTCA TGAAAACAGG GTAAAGGACA TGTTCAGTTT  840
TGAAATTCCA GAGCATCATG TTATTGTTTC AGCATCGAGT GATGGTTTCA TCAAAATGTG  900
GAAGCTTAAG CAGGATAAGA AAGTTCCCCC ATCTTTACTC TGTGAAATAA ACACTAATGC  960
CAGGCTGACG TGTCTTGGAG TGTGGCTAGA CAAAGTGGCA GACATGAAAA GCCTTCCCTC 1020
AGCTGCAGAG CCTTCTCCTG TAAGTAAAGA ACAGTCCAAA ATTGGCAAAA AGGAGCCTGG 1080
TGACACAGTG CACAAAGAAG AAAAGCGGTC AAAACCTAAC ACAAAGAAAC GCGGTTTAAC 1140
```

```
AGGTGACAGT AAGAAAGCAA CAAAAGAAAG TGGCCTGATA TCAACCAAGA AGAGGAAAAT 1200
GGTAGAAATG TTGGAAAAGA AGAGGAAAAA GAAGAAAATA AAAACAATGC AGTGAATCAC 1260
AGATGTCTCC TGAAAGAACT CTTTTAGATG AAATCATTCT ACTCAAATGT ACCTTAATTT 1320
TTTTTTTTCC CTGAGTAAAA GCAAGAAATT TCTTCCTTTG GAAAAAATAT ATATATTAAA 1380
AAACCACTTT TAGATGGTTT TTTTTAAAAA AAAAAAAAAA ACTGGTAAAA TTACTTTTGG 1440
CAGACAGTGT TTTATGAATT ATGTATCATG TTGATATATA ATATGTTAAT GTGTCATGTA 1500
ATTTTTACTT TGTACAAAGC AAATAAAGAT CTTTCTCAAA AAAAAAAAAA AAAA
```

SEQ ID NO:152 PFG1 Protein sequence:
Protein Accession #:   NP_060376.1

```
1    11   21   31   41   51
|    |    |    |    |    |
MELVAGCYEQ VLFGFAVHPE PKACGDHEQW TLVADFTHHA HTASLSAVAV NSRFVVTGSK   60
DETIHIYDMK KKIEHGALVH HSGTITCLKF YGNRHLISGA EDGLICIWDA KKWECLKSIK  120
AHKGQVTFLS IHPSGKLALS VGTDKTLRTW NLVEGRSAFI KNIKQNAHIV EWSPRGEQYV  180
VIIQNKIDIY QLDTASISGT ITNEKRISSV KFLSESVLAV AGDEEVIRFF DCDSLVCLCE  240
FKAHENRVKD MFSFEIPEHH VIVSASSDGF IKMWKLKQDK KVPPSLLCEI NTNARLTCLG  300
VWLDKVADMK SLPPAAEPSP VSKEQSKIGK KEPGDTVHKE EKRSKPNTKK RGLTGDSKKA  360
TKESGLISTK KRKMVEMLEK KRKKKKIKTM Q
```

SEQ ID NO:153 PFD6 DNA SEQUENCE

Nucleic Acid Accession #:   NM_014668
Coding sequence: 110-2953   (underlined sequences correspond to start and stop codons)

```
1    11   21   31   41   51
|    |    |    |    |    |
GATGTCTTGG ACATGCTCTG GCTGGCTAAT CTCCATGTTC TAGCCGACTG AAAAATACGGT   60
GGCCAAGTGG ATGGTGTGCT TATTTGCAGT CTAAAGAAAT TTCCTTTTGA TGTGGCAGAA  120
AATCGAGGAT GTGGAGTGGA GACCCCAGAC TTACTTGGAG CTGGAGGGTC TGCCTTGCAT  180
CCTGATCTTC AGTGGGATGG ACCCGCATGG GGAGTCCTTG CCGAGGTCTT TGAGGTACTG  240
TGACCTGCGA TTGATAAACT CCTCCTGCTT GGTGAGAACA GCCTTGGAGC AGGAGCTGGG  300
CCTGGCTGCC TACTTTGTGA GCAACGAGGT TCCCTTGGAG AAGGGGGGCTA GGAACGAGGC  360
CTTGGAGAGT GATGCTGAGA AGCTGAGCAG CACAGACAAC GAGGATGAGG AGCTGGGGAC  420
AGAAGGCTCT ACCTCGGAGA AGAGAAGCCC CATGAAGAGG TCCCGCTCCCACGA  480
CTCAGCATCC TCATCCCTCT CCTCCAAGGC TTCCGGTTCA GCGCTCGGTG GCGAGTCCTC  540
GGCTCAGCCC ACAGCACTCC CCCAGGGAGA GCATGCCAGG TCGCCCCAGC CCCGTGGCCC  600
CGCAGAGGAG GGCAGAGCCC CTGGTGAGAA ACAGAGGCCC CGGGCAAGTC AGGGGCCACC  660
CTCGGCCATC AGCAGGCACA GTCCCGGGCC GACGCCCACG CCCGACTGTA GCCTCAGGAC  720
CGGCCAGAGG AGCGTCCAGG TGTCGGTCAC CTCGTCGTGC TCCCAGCTGT CCTCCTCCTC  780
GGGCTCATCC TCCTCATCCG TGGCGCCCGC TGCCGGCACG TGGGTCCTGC AGGCCTCCCA  840
GTGCTCCTTG ACCAAGGCCT GCCGCCAGCC ACCCATTGTC TTCTTGCCCA AGCTCGTGTA  900
CGACATGGTT GTGTCCACTG ACAGCAGTGG CCTGCCCAAG GCCGCCTCCC TCCTGCCCTC  960
CCCCTCGGTC ATGTGGGCCA GCTCTTTCCG CCCCCTGCTC AGCAAGACCA TGACATCCAC 1020
CGAGCAGTCC CTCTACTACC GGCAGTGGAC GGTGCCCCGG CCCAGCCACA TGGACTACGG 1080
CAACCGGGCC GAGGGCCGCG TGGACGGCTT CCACCCCCGC AGGCTGCTGC TCAGCGGCCC 1140
CCCTCAGATC GGGAAGACAG GTGCCTACCT GCAGTTCCTC AGTGTCCTGT CCAGGATGCT 1200
TGTTCGGCTC ACAGAAGTGG ATGTCTATGA CGAGGAGGAG ATCAATATCA ACCTCAGAGA 1260
AGAATCTGAC TGGCATTATC TCCAGCTTAG CGACCCCTGG CCAGACCTGG AGCTGTTCAA 1320
GAAGTTGCCC TTTGACTACA TCATTCACGA CCCGAAGTAT GAAGATGCCA GCCTGATTTG 1380
TTCGCACTAT CAGGGTATAA AGAGTGAAGA CAGAGGGATG TCCCGGAAGC CGGAGGACCT 1440
TTATGTGCGG CGTCAGACGG CACGGATGAG ACTGTCCAAG TACGCAGCGT ACAACACTTA 1500
CCACCACTGT GAGCAGTGCC ACCAGTACAT GGGCTTCACC ACCCCGCTACC AGCTGTATGA 1560
GTCCACCCTG CACGCCTTTG CCTTCTCTTA CTCCATGCTA GGAGAGGGAGA TCCAGCTGCA 1620
CTTCATCATC CCCAAGTCCA AGGAGCACCA CTTTGTCTTC AGCCAACCTG GAGGCCAGCT 1680
GGAGACGCATG CGACTACCCC TCGTGACAGA CAAGAGCCAT GAATATATAA AAAGTCCGAT 1740
ATTCACTCCA ACCACCGGCC GTCACGAACA TGGGCTCTTT AATCTGTACC ACGCAATGGA 1800
CGGTGCCAGC CATTTGCACG TGCTGGTTGT CAAGGAATAC GAGATGGCAA TTTATAAGAA 1860
ATATTGGCCC AACCACATCA TGCTGGTGCT CCCCAGTATC TTCAACAGTG CTGGAGTTGG 1920
TGCTGCTCAT TTCCTCATCA AGGAGCTGTC CTACCATAAC CTGGAGCTCG AGCGGAACCG 1980
GCAGGAGGAG CTGGGAATCA AGCCGCAGGA CATCTGGCCT TTCATTGTGA TCTCTGATGA 2040
CTCCTGCGTG ATGTGGAACG TGGTGGATGT CAACTCTGCT GGGGAGAGAA GCAGGGAGTT 2100
CTCCTGGTCG GAAAGGAACG TGTCTTTGAA GCACATCATG CAGCCATCG AGGCGGCCCC 2160
CGACATCATG CACTACGCCC TGCTGGGCCT GCGGAAGTGG TCCAGCAAGA CCCGGGCCAG 2220
CGAGGTGCAA GAGCCCTTCT CCCGCTGCCA CGTGCACAAC TTCATCATCC TGAACGTGGA 2280
CCTGACCCAG AACGTGCAGT ACAACCAGAA CCGGTTCCTG TGTGACGATG TAGACTTCAA 2340
CCTGCGGGTG CACAGCGCCG GCCTCCTGCT CTGCCGGTTC AACCGCTTCA GCGTGATGAA 2400
GAAGCAGATC GTGGTGGGCG GCCACAGGTC CTTCCACATC ACATCCAAGG TGTCTGATAA 2460
CTCTGCCGCG GTCGTGCCGG CCCAGTACAT CTGTGCCCCG GACAGCAAGC ACACGTTCCT 2520
CGCAGCGCCC GCCCAGCTCC TGCTGGAGAA GTTCCTGCAG CACCACAGCC ACCTCTTCTT 2580
CCCCGCTGTCC CTGAAGAACC ATGACCACCC AGTGCTGTCT GTCGACTGTT ACCTGAACCT 2640
GGGATCTCAG ATTTCTGTTT GCTATGTGAG CTCCAGGCCC CACTCTTTAA ACATCAGCTG 2700
CTCGGACTTG CTGTTCAGTG GGCTGCTGCT GTACCTCTGT GACTCTTTTG TGGGAGCTAG 2760
CTTTTTGAAA AAGTTTCATT TTCTGAAAGG TGCGACGTTG TGTGTCATCT GTCAGGACCG 2820
GAGCTCACTG CGCCAGACGG TCGTCCGCCT GGAGCTCGAG GACGAGTGGC AGTTCCGGCT 2880
```

```
GCGCGATGAG TTCCAGACCG CCAATGCCAG GGAAGACCGG CCGCTCTTTT TTCTGACGGG 2940
ACGACACATC TGAGGAAGAC AGCGGCGAGT TTTCTGAAGA GATGAGTGCT CAGAGCCCTC 3000
ATGCTGTTGA GGCTAAAGGG AGGCCTGGAA CGGTGGGGCG TTTGACTGGA ATGGACCCCA 3060
GGGACTGTCC AGGTGCAGCC CCTCCTAGTA CACATGGGCC CCCGAGGCCG TGGTCCTGGG 3120
AGCCAGGAAG ACTCCGCAGT GGGTGAGAAT GAAAACTTGA GACTCCCAAG TTCTGGGCCA 3180
GCCCATTGCT CTGGGCTGTT TTAAAGCCCA TTTCACGAGG AACAAAGATT TACTTCCTGT 3240
CCTGCCATTC GTGTGCTTCC ATGGACAAAC CTGATTTTTT TCTCTTAGTT CTAAAGAATC 3300
TTGGGTTATT TTGTAGCGGT GCCAGTATTT CAGTAGATGG GATTTCAGCC AAGTAGGTTC 3360
CCCTGTAACC TCCTACAAAG CAATATTCCA AAGGAACATT TTAACTGTAA AGGCTGGAGA 3420
CAAGAAAAAA TAAGTAGATC GTTTTAATAA CAATTATTTA ATTGCCTATA AGTTTGCTGT 3480
TTCAGAGGCT AGCCCAAAGG CATCAAATTT AATAAAGTTA AACAAATTGA TTTACTTCAG 3540
AGCAAATATG ATCCTATTAA AATAATATAG GGTAAATACC CTACCTCTTA GAAAGGGCAA 3600
AAATGCAAAG AAGCTTTCTT TAAAACTAAA AGGGTTTTTT GGGGGGGGAG TTGGCGGGGA 3660
GGAAATAAGG CTAACAGAGG TTGACCTAAA ATTAGCCTTA CAAAGGAGAA AGGACCACAT 3720
TGCTTACTTG AAACAGACAA TGAAAACAAC CAAAGTGATA TATAAAATAG TTGATGAGAA 3780
CTAGACTTAT GACTGTAGTT TACTAGAGTT TAGTTTTCAG TTGCTGAAGT AGCTCATTTT 3840
CTCTTACTAA TGTTTGGTTC CTCAGGGAAG AATCTCACTT GACTAGAGAG GAGGTGGGAA 3900
CAGAAGAGAG AAGGAGGCAG GGAGATGTAT TTCTTAGGGC TCACCCCTTC ACAGACTGAC 3960
AGAATGGTTT TGTTTTGTTT TGTTTTGTTT TGTTTTGTTT TTGAGATGGA CTCTAGCTCT 4020
GTCACCCAGG CTGGAGTGCA GTGGTGCGAT CTCGGCTCAC TGCAAGCTCC GCCTCCCGGG 4080
TTCTCACCAT TCTCCTGCCT CAGCGTCCCG AGTAGCTGGG ACTACAGGCG CCCACCACCA 4140
CGCCCGGCTA ATTTTTTGTA TTTTTTAGTA GAGACGGGGT TTCACCATGT TAGCCAGGAT 4200
GGTCTCGATC TCCTGACCTC GTGATCCGCC CGCCTCGGCC TCCCAAAGTG CTGGGATTAC 4260
AGGCGTGAGC CACCGTGCCT GCCCCAGAAT GGTTTTTAAA GCCACAGTTG AGAGGCCACC 4320
CATTGCCCGG CGCCTGGACA GTGATCATCT TGTTCATCTT GTTCAGTCCT TTCTTGTGTG 4380
ATTGGAATTA TTCATCCCCT TTGAAAGATG AGAAGGTTGA GATGCAAAGA GTCTACCTTT 4440
CCAAGTTCTC ACTGCTGGAA AGAGCTAGAA GCACAGTTCA AAGTTCTGGC TTCTGGACTC 4500
TGCAGTCCAG GTCTCCCTTC TCCCACTTGC CTACCCTCAA TGCCACACTG TTTTTGAAGT 4560
GGCCCATAAC TTGAAGGAAA AGTTTAAAGA CAGTTCAATT TAATCATCAG AATGCATTCT 4620
TTTTTTTTTC GGAGACGGAG TTTCACTCTT GCTGCCCAGG CTGGAGTGCA ATGGTGCAAT 4680
GATCTCGGCT CACTGCAACC TCTGCCTCCT GGGTTCAAGT GATTCTCCAG CCTCAGCCTC 4740
CCGAGTAGCT GGGATTATGG GCGCCCACCA CCATGCCCAG CTAATTTTTG TATTTTTTTT 4800
TTTTAGTAGA GATGGGGTTT CGCCAGGTTG GCCAGGCTGG TCTTGTGAAC TCCTGGCCTC 4860
AGGTGATCTG CCCACCTCAT CCTCCAAAAG TGCTGGGATT ACAGGCATGA GCCACTGCGC 4920
CTGGCCTCGA AATGCATTCT TACACATCTA TCCTAGACAT TTATAAGCAC TCTAATGGAT 4980
AACAATCCAA GAATAAATGA TTGTAAAAGA TGATGCCGAA GAGTTGATGT CAATCTTTTT 5040
TTCCTAAGAA AAAAAGTCCG CGAGTATTAA ATATTTAGAT CAATGTTTAT AAAATGATTA 5100
CTTTGTATAT CTCATTATTC CTATTTTGGA ATAAAAACTG ACCTTCTTTA ATCATATACT 5160
TGTCTTTTGT AAATAGCAGC TTTTGTGTCA TTCTCCCCAC TTTATTAGTT AATTTAAATT 5220
GGAAAAAACC CTCAAACTAA TATTCTTGTC TGTTCCAGTC TTATAAATAA AACTTATAAT 5280
GCATG
```

SEQ ID NO:154 PFD6 Protein sequence:
Protein Accession #:   NP_055483.1

```
1    11    21    31    41    51
|    |     |     |     |     |
MWQKIEDVEW RPQTYLELEG LPCILIFSGM DPHGESLPRS LRYCDLRLIN SSCLVRTALE    60
QELGLAAYFV SNEVPLEKGA RNEALESDAE KLSSTDNEDE ELGTEGSTSE KRSPMKRERS   120
RSHDSASSSL SSKASGSALG GESSAQPTAL PQGEHARSPQ PRGPAEEGRA PGEKQRPRAS   180
QGPPSAISRH SPGPTPQPDC SLRTGQRSVQ VSVTSSCSQL SSSSGSSSSS VAPAAGTWVL   240
QASQCSLTKA CRQPPIVFLP KLVYDMVVST DSSGLPKAAS LLPSPSVMWA SSFRPLLSKT   300
MTSTEQSLYY RQWTVPRPSH MDYGNRAEGR VDGFHPRRLL LSGPPQIGKT GAYLQFLSVL   360
SRMLVRLTEV DVYDEEEINI NLREESDWHY LQLSDPWPDL ELFKKLPFDY IIHDPKYEDA   420
SLICSHYQGI KSEDRGMSRK PEDLYVRRQT ARMRLSKYAA YNTYHHCEQC HQYMGFHPRY   480
QLYESTLHAF AFSYSMLGEE IQLHFIIPKS KEHHFVFSQP GGQLESMRLP LVTDKSHEYI   540
KSPTFIPTTG RHEHGLFNLY HAMDGASHLH VLVVKEYEMA IYKKYWPNHI MLVLPSIFNS   600
AGVGAAHFLI KELSYHNLEL ERNRQEELGI KPQDIWPFIV ISDDSCVMWN VVDVNSAGER   660
SREFSWSERN VSLKHIMQHI EAAPDIMHYA LLGLRKWSSK TRASEVQEPF SRCHVHNFII   720
LNVDLTQNVQ YNQNRFLCDD VDFNLRVHSA GLLLCRFNRF SVMKKQIVVG GHRSFHITSK   780
VSDNSAAVVP AQYICAPDSK HTFLAAPAQL LLEKFLQHHS HLFFPLSLKN HDHPVLSVDC   840
YLNLGSQISV CYVSSRPHSL NISCSDLLFS GLLLYLCDSF VGASFLKKFH FLKGATLCVI   900
CQDRSSLRQT VVRLELEDEW QFRLRDEFQT ANAREDRPLF FLTGRHI
```

SEQ ID NO:155 PFC6 DNA SEQUENCE

Nucleic Acid Accession #: NM_000522
Coding sequence: 1-1167 (underlined sequences correspond to start and stop codons)

```
1    11    21    31    41    51
|    |     |     |     |     |
ATGACAGCCT CCGTGCTCCT CCACCCCCGC TGGATCGAGC CCACCGTCAT GTTTCTCTAC    60
GACAACGGCG GCGGCCTGGT GGCCGACGAG CTCAACAAGA ACATGGAAGG GGCGGCGGCG   120
GCTGCAGCAG CGGCTGCAGC GGCGGCCGGCT GCCGGGGCCG GGGGCGGGGG CTTCCCCCAC   180
CCGGCGGCTG CGGCGGCAGG GGGCAACTTC TCGGTGGCGG CCGCGGCCGC GGCTGCGGCG   240
GCCGCCGCGG CCAACCAGTG CCGCAACCTG ATGGCGCACC CGGCGCCCTT GGCGCCAGGA   300
GCCGCGTCCG CCTACAGCAG CGCCCCCGGG GAGGCGCCCC CGTCGGCTGC CGCCGCTGCT   360
```

361

GCCGCGGCTG CCGCTGCAGC CGCCGCCGCC GCCGCCGCGT CGTCCTCGGG AGGTCCCGGC 420
CCGGCGGGCC CGGCGGCGGC AGAGGCGGCC AAGCAATGCA GCCCCTGCTC GGCAGCGGCG 480
CAGAGCTCGT CGGGGCCCGC GGCGCTGCCC TATGGCTACT TCGGCAGCGG CTACTACCCG 540
TGCGCCCGCA TGGGGCCCGCC CCCCAACGCC ATCAAGTCGT GCCCCCAGCC CCCCTCGGCC 600
GCCGCCGCCG CCGCCTTCGC GGACAAGTAC ATGGATACCG CCGGCCCAGC TGCCGAGGAG 660
TTCAGCTCCC GCGCTAAGGA GTTCGCGTTC TACCACCAGG GCTACGCAGC CGGGCCTTAC 720
CACCACCATC AGCCCATGCC TGGCTACCTG GATATGCCAG TGGTGCCGGG CCTCGGGGGC 780
CCCGGCGAGT CGCGCCACGA ACCCTTGGGT CTTCCCATGG AAAGCTACCA GCCCTGGGCG 840
CTGCCCAACG GCTGGAACGG CCAAATGTAC TGCCCCAAAG AGCAGGCGCA GCCTCCCCAC 900
CTCTGGAAGT CCACTCTGCC CGACGTGGTC TCCCATCCCT CGGATGCCAG CTCCTATAGG 960
AGGGGGAGAA AGAAGCGCGT GCCTTATACC AAGGTGCAAT TAAAAGAACT TGAACGGGAA 1020
TACGCCACGA ATAAATTCAT TACTAAGGAC AAACGGAGGC GGATATCAGC CACGACGAAT 1080
CTCTCTGAGC GGCAGGTCAC AATCTGGTTC CAGAACAGGA GGGTTAAAGA GAAAAAAGTC 1140
ATCAACAAAC TGAAAACCAC TAGT<u>TAA</u>

SEQ ID NO:156 <u>PFC6 Protein sequence:</u>
Protein Accession #:   NP_000513.1

```
1      11     21     31     41     51
|      |      |      |      |      |
MTASVLLHPR WIEPTVMFLY DNGGGLVADE LNKNMEGAAA AAAAAAAAAA AGAGGGGFPH   60
PAAAAAGGNF SVAAAAAAAA AAAANQCRNL MAHPAPLAPG AASAYSSAPG EAPPSAAAAA  120
AAAAAAAAAA AAASSSGGPG PAGPAAAEAA KQCSPCSAAA QSSSGPAALP YGYFGSGYYP  180
CARMGPPPNA IKSCPQPPSA AAAAAFADKY MDTAGPAAEE FSSRAKEFAF YHQGYAAGPY  240
HHHQPMPGYL DMPVVPGLGG PGESRHEPLG LPMESYQPWA LPNGWNGQMY CPKEQAQPPH  300
LWKSTLPDVV SHPSDASSYR RGRKKRVPYT KVQLKELERE YATNKFITKD KRRRISATTN  360
LSERQVTIWF QNRRVKEKKV INKLKTTS
```

SEQ ID NO:157 PFA3 DNA SEQUENCE
Nucleic Acid Accession #:   AW102723
Coding sequence: 523-2676 (underlined sequences correspond to start and stop codons)

```
1      11     21     31     41     51
|      |      |      |      |      |
CCCTTATGGC GATTGGGCGG CTGCAGAGAC CAGGACTCAG TTCCCCTGCC CTAGTCTGAG   60
CCTAGTGGGT GGGACTCAGC TCAGAGTCAG TTTTCAGAAG CAGGTTTCAG TTGCAGAGTT  120
TTCCTACACT TTTCCTGCGC TAGAGCAGCG AGCAGCCTGG AACAGACCCA GGCGGAGGAC  180
ACCTGTGGGG GAGGGAGCGC CTGGAGGAGC TTAGAGACCC CAGCCGGGCG TGATCTCACC  240
ATGTGCGGAT TTGCGAGGCG CGCCCTGGAG CTGCTAGAGA TCCGGAAGCA CAGCCCCGAG  300
GTGTGCGAAG CCACCAAGAC TGCCGGCTCTT GGAGAAAGCG TGAGCAGGGG GCCACCGCGG  360
TCTCCGGCCT GTCTGCACCC TGTCGCCTGA GCTGCCTGAC AGTGACAATG ACATCCCAGT  420
TACCAGTGTC CTTGAATTGA TAGTGGCTTC TGTTTGTCAG TCTCATATAA GAACTACAGC  480
TCATCAGGAG GAGATCGCAG CAGGGTAAGA GACACCAACA CC<u>ATG</u>TTCTG CACGAAGCTC  540
AAGGATCTCA AGATCACAGG AGAGTGTCCT TTCTCCTTAC TGGCACCAGG TCAAGTTCCT  600
AACGAGTCTT CAGAGGAGGC AGCAGGAAGC TCAGAGAGCT GCAAAGCAAC CGTGCCCATC  660
TGTCAAGACA TTCCTGAGAA GAACATACAA GAAAGTCTTC CTCAAAGAAA AACCAGTCGG  720
AGCCGAGTCT ATCTTCACAC TTTGGCAGAG AGTATTTGCA AACTGATTTT CCCAGAGTTT  780
GAACGGCTGA ATGTTGCACT TCAGAGAACA TTGGCAAAGC ACAAAATAAA AGAAAGCAGG  840
AAATCTTTGG AAAGAGAAGA CTTTGAAAAA ACAATTGCAG AGCAAGCAGT GCAGCAGAGT  900
CCAGTGGAGT TATCAAAGAA TCTCTTGGTG AAGAGGTTTT TAAAATATGT TACGAGGAAG  960
ATGAAAACAT CCTTGGGGTG GTTGGAGGCA CCCTTAAAGA TTTTTAAACA GCTTCAGTAC  1020
CCTTCTGAAA CAGAGCAGCC ATTGCCAAGA AGCAGGAAAA AGGGGCAGCT TGAGGACGCC  1080
TCCATTCTAT GCCTGGATAA GGAGGATGAT TTTCTACATG TTTACTACTT CTTCCCTAAG  1140
AGAACCACCT CCCTGATTCT TCCCGGCATC ATAAAGGCAG CTGCTCACGT ATTATATGAA  1200
ACGGAAGTGG AAGTGTCGTT AATGCCTCCC TGCTTCCATA ATGATTGCAG CGAGTTTGTG  1260
AATCAGCCCT ACTTGTTGTA CTCCGTTCAC ATGAAAAGCA CCAAGCCATC CCTGTCCCCC  1320
AGCAAACCCC AGTCCTCGCT GGTGATTCCC ACATCGCTAT TCTGCAAGAC ATTTCCATTC  1380
CATTTCATGT TTGACAAAGA TATGACAATT CTGCAATTTG GCAATGGCAT CAGAAGGCTG  1440
ATGAACAGGA GAGACTTTCA AGGAAAGCCT AATTTTGAAT ACTTTGAAAT TCTGACTCCA  1500
AAAATCAACC AGACCTTTAG CGGGATCATG ACTATGTTGA ATATGCAGTT TGTTGTACGA  1560
GTGAGGAGAT GGGACAACTC TGTGAAGAAA TCTTCAAGGG TTATGGACCT CAAAGGCCAA  1620
ATGATCTACA TTGTTGAATC CAGTGCAATC TTGTTTTTGG GGTCACCCTG TGTGGACAGA  1680
TTAGAAGATT TTACAGGACG AGGGCTCTAC CTCTCAGACA TCCCAATTCA CAATGCACTG  1740
AGGGATGTGG TCTTAATAGG GGAACAAGCC CGAGCTCAAG ATGGCCTGAA GAAGAGGCTG  1800
GGGAAGCTGA AGGCTACCCT TGAGCAAGCC CACCAAGCCC TGGAGGAGGA GAAGAAAAAG  1860
ACAGTAGACC TTCTGTGCTC CATATTTCCC TGTGAGGTTG CTCAGCAGCT GTGGCAAGGG  1920
CAAGTTGTGC AAGCCAAGAA GTTCAGTAAT GTCACCATGC TCTTCTCAGA CATCGTTGGG  1980
TTCACTGCCA TCTGCTCCCA GTGCTCACCG CTGCAGGTCA TCACCATGCT CAATGCACTG  2040
TACACTCGCT TCGACCAGCA GTGTGGAGAG CTGGATGTCT ACAAGGTGGA GACCATTGCG  2100
ATGCCTATTG TGTGGCTTGG GGGATTACAC AAAGAGAGTG ATACTCATGC TGTTCAGATA  2160
GCGCTGATGG CCCTGAAGAT GATGGAGCTC TCTGATGAAG TTATGTCTCC CCATGGAGAA  2220
CCTATCAAGA TGCGAATTGG ACTGCACTCT GGATCAGTTT TTGCTGGCGT CGTTGGAGTT  2280
AAAATGCCCC GTTACTGTCT TTTTGGAAAC AATGTCACTC TGGCTAACAA ATTTGAGTCC  2340
TGCAGTGTAC CACGAAAAAAT CAATGTCAGC CCAACAACTT ACAGATTACT CAAAGACTGT  2400
CCTGGTTTCG TGTTTACCCC TCGATCAAGG GAGGAACTTC CACCAAACTT CCCTAGTGAA  2460
ATCCCCGGAA TCTGCCATTT TCTGGATGCT TACCAACAAG GAACAAACTC AAAACCATGC  2520
```

```
TTCCAAAAGA AAGATGTGGA AGATGCAAGC CAATTTTTTA GGCAAAGCAT CAGGAATAGA 2580
TTAGCAACCT ATATACCTAT TTATAAGTCT TTGGGGTTTG ACTCATTGAA GATGTGTAGA 2640
GCCTCTGAAA GCACTTTAGG GATTGTGAT GGCTAACAAG CAGTATTAAA ATTTCAGGAG 2700
CCAAGTCACA ATCTTTCTCC TGTTTAACAT GACAAAATGT ACTCACTTCA GTACTTCAGC 2760
TCTTCAAGAA AAAAAAAAAA ACCTTAAAAA GCTACTTTTG TGGGAGTATT TCTATTATAT 2820
AACCAGCACT TACTACCTGT ACTCAAAATT CAGCACCTTG TACATATATC AGATAATTGT 2880
AGTCAATTGT ACAAACTGAT GGAGTCACCT GCAATCTCAT ATCCTGGTGG AATGCCATGG 2940
TTATTAAAGT GTGTTTGTGA TAGTTGTCGT CAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 3000
AAAA
```

SEQ ID NO:158 PFA3 Protein sequence:
Protein Accession #:   NP_000847.1

```
1    11   21   31   41   51
|    |    |    |    |    |
MFCTKLKDLK ITGECPFSLL APGQVPNESS EEAAGSSESC KATVPICQDI PEKNIQESLP  60
QRKTSRSRVY LHTLAESICK LIFPEFERLN VALQRTLAKH KIKESRKSLE REDFEKTIAE  120
QAVQQSPVEL SKNLLVKRFL KYVTRKMKTS LGWLEAPLKI FKQLQYPSET EQPLPRSRKK  180
GQLEDASILC LDKEDDFLHV YYFFPKRTTS LILPGIIKAA AHVLYETEVE VSLMPPCFHN  240
DCSEFVNQPY LLYSVHMKST KPSLSPSKPQ SSLVIPTSLF CKTFPFHFMF DKDMTILQFG  300
NGIRRLMNRR DFQGKPNFEY FEILTPKINQ TFSGIMTMLN MQFVVRVRRW DNSVKKSSRV  360
MDLKGQMIYI VESSAILFLG SPCVDRLEDF TGRGLYLSDI PIHNALRDVV LIGEQARAQD  420
GLKKRLGKLK ATLEQAHQAL EEEKKKTVDL LCSIFPCEVA QQLWQGQVVQ AKKFSNVTML  480
FSDIVGFTAI CSQCSPLQVI TMLNALYTRF DQQCGELDVY KVETIAMPIV WLGGLHKESD  540
THAVQIALMA LKMMMELSDEV MSPHGEPIKM RIGLHSGSVF AGVVGVKMPR YCLFGNNVTL  600
ANKFESCSVP RKINVSPTTY RLLKDCPGFV FTPRSREELP PNFPSEIPGI CHFLDAYQQG  660
TNSKPCFQKK DVEDASQFFR QSIRNRLATY IPIYKSLGFD SLKMCRASES TLGIVDG
```

SEQ ID NO:159 PFA1 DNA SEQUENCE
Nucleic Acid Accession #:   NM_004362
Coding sequence: 102-1934 (underlined sequences correspond to start and stop codons)

```
1    11   21   31   41   51
|    |    |    |    |    |
CGCCGGCGGG ACTGGTCTGA AGAGACGCGG GGACAAAGTG GCAACGACTT GGACATCTGA  60
GCTGTCACTG CCGAAAACAG GCCGCAAGAG AGATAACTGA TATGCATTTC CAAGCCTTTT  120
GGCTATGTTT GGGTCTTCTG TTCATCTCAA TTAATGCAGA ATTTATGGAT GATGATGTTG  180
AGACGGAAGA CTTTGAAGAA AATTCAGAAG AAATTGATGT TAATGAAAGT GAACTTTCCT  240
CAGAGATTAA ATATAAGACA CCTCAACCTA TAGGAGAAGT ATATTTTGCA GAAACTTTTG  300
ATAGTGGAAG GTTGGCTGGA TGGGTCTTAT CAAAAGCAAA GAAAGATGAC ATGGATGAGG  360
AAATTTCAAT TACGATGGA AGATGGGAAA TTGAAGAGTT GAAAGAAAAC CAGGTACCTG  420
GTGACAGAGG ACTGGTATTA AAATCTAGAG CAAAGCATCA TGCAATATCT GCTGTATTAG  480
CAAAACCATT CATTTTTGCT GATAAACCCT TGATAGTTCA ATATGAAGTA AATTTTCAAG  540
ATGGTATTGA TTGTGGAGGT GCATACATTA AACTCCTAGC AGACACTGAT GATTTGATTC  600
TGGAAAACTT TTATGATAAA ACATCCTATA TCATTATGTT TGGACCAGAT AAATGTGGAG  660
AAGATTATAA ACTTCATTTT ATCTTCAGAC ATAAACATCC CAAAACTGGA GTTTTCGAAG  720
AGAAACATGC CAAACCTCCA GATGTAGACC TTAAAAAGTT CTTTACAGAC AGGAAGACTC  780
ATCTTTATAC CCTTGTGATG AATCCAGATG ACACATTTGA GGTGTTAGTT GATCAAACAG  840
TTGTAAACAA AGGAAGCCTC CTAGAGGATG TGGTTCCTCC TATCAAACCT CCCAAAGAAA  900
TTGAAGATCC CAATGATAAA AAACCTGAGG AATGGGATGA AAGAGCAAAA ATTCCTGATC  960
CTTCTGCCGT CAAACCAGAA GACTGGGATG AAAGTGAACC TGCCCAAATA GAAGATTCAA  1020
GTGTTGTTAA ACCTGCTGGC TGGCTTGATG ATGAACCAAA ATTTATCCCT GATCCTAATG  1080
CTGAAAAACC TGATGACTGG AATGAAGACA CGGATGGGCA ATGGGAGGCA CCTCAGATTC  1140
TTAATCCAGC ATGTCGGATT GGGTGTGGTA AGTGGAAACC TCCCATGATA GATAACCCAA  1200
AATACAAAGG AGTATGGAGA CCTCCACTGG TCGATAATCC TAACTATCAG GGAATCTGGA  1260
GTCCTCGAAA AATTCCTAAT CCAGATTATT TCGAAGATGA TCATCCATTT CTTCTGACTT  1320
CTTTCAGTGC TCTTGGTTTA GAGCTTTGGT CTATGACCTC TGATATCTAC TTTGATAATT  1380
TTATTATCTG TTCGGAAAAG GAAGTAGCAG ATCACTGGGC TGCAGATGGT TGGAGATGGA  1440
AAATAATGAT AGCAAATGCT AATAAGCCTG GTGTATTAAA ACAGTTAATG GCAGCTGCTG  1500
AAGGGCACCC ATGGCTTTGG TTGATTTATC TTGTGACAGC AGGAGTGCCA ATAGCATTAA  1560
TTACTTCATT TTGTTGGCCA AGAAAAGTAA AGAAAAAACA TAAAGATACA GAGTATAAAA  1620
AAACCGACAT ATGTATACCA CAAACAAAAG GAGTACTAGA GCAAGAAGAA AAGGAAGAGA  1680
AAGCAGCCCT GGAAAAACCA ATGGACCTGG AAGAGGAAAA AAAGCAAAAT GATGGTGAAA  1740
TGCTTGAAAA AGAAGAGGAA AGTGAACCTG AGGAAAAGAG TGAAGAAGAA ATTGAAATCA  1800
TAGAAGGGCA AGAAGAAAGT AATCAATCAA ATAAGTCTGG GTCAGAGGAT GAGATGAAAG  1860
AAGCAGATGA GAGCACAGGA TCTGGAGATG GGCCGATAAA GTCAGTACGC AAAAGAAGAG  1920
TACGAAAGGA CTAAACTAGA TTGAAATATT TTTAATTCCC GAGAGGATGT TTGGCATTGT  1980
AAAAATCAGC ATGCCAGACC TGAACTTTAA TCAGTCTGCA CATCCTGTTT CTAATATCTA  2040
GCAACATTAT ATTCTTTCAG ACATTTATTT TAGTCCTTCA TTTCCGAGGA AAAAGAAGCA  2100
ACTTTGAAGT TACCTCATCT TTGAATTTAG AATAAAAGTG GCACATTACA TATCGGATCT  2160
AAGAGATTAA TACCATTAGA AGTTACACAG TTTTAGTTGT TTGGAGATAG TTTTGGTTTG  2220
TACAGAACAA AATAATATGT AGCAGCTTCA TTGCTATTGG AAAAATCAGT TATTGGAATT  2280
TCCACTTAAA TGGCTATACA ACAATATAAC TGGTAGTTCT ATAATAAAAA TGAGCATATG  2340
TTCTGTTGTG AAGAGCTAAA TGCAATAAAG TTTCTGTATG GTTGTTTGAT TCTATCAACA  2400
ATTGAAAGTT TTGTATATGA CCCACATTTA CCTAGTTTGT GTCAAATTAT AGTTACAGTG  2460
AGTTGTTTGC TTAAATTATA GATTCCTTTA AGGACATGCC TTGTTCATAA AATCACTGGA  2520
```

TTATATTGCA GCATATTTTA CATTTGAATA CAAGGATAAT GGGTTTTATC AAAACAAAAT 2580
GATGTACAGA TTTTTTTTCA AGTTTTTATA GTTGCTTTAT GCCAGAGTGG TTTACCCCAT 2640
TCACAAAATT TCTTATGCAT ACATTGCTAT TGAAAATAAA ATTTAAATAT TTTTTCATCC 2700

5   TGAAAAAAAA

SEQ ID NO:160 PFA1 Protein sequence:
Protein Accession #:   NP_004353.1

10   1   11   21   31   41   51
       | | | | | |
       MHFQAFWLCL GLLFISINAE FMDDDVETED FEENSEEIDV NESELSSEIK YKTPQPIGEV   60
       YFAETFDSGR LAGWVLSKAK KDDMDEEISI YDGRWEIEEL KENQVPGDRG LVLKSRAKHH 120
15   AISAVLAKPF IFADKPLIVQ YEVNFQDGID CGGAYIKLLA DTDDLILENF YDKTSYIIMF 180
       GPDKCGEDYK LHFIFRHKHP KTGVFEEKHA KPPDVDLKKF FTDRKTHLYT LVMNPDDTFE 240
       VLVDQTVVNK GSLLEDVVPP IKPPKEIEDP NDKKPEEWDE RAKIPDPSAV KPEDWDESEP 300
       AQIEDSSVVK PAGWLDDEPK FIPDPNAEKP DDWNEDTDGE WEAPQILNPA CRIGCGEWKP 360
       PMIDNPKYKG VWRPPLVDNP NYQGIWSPRK IPNPDYFEDD HPFLLTSFSA LGLELWSMTS 420
20   DIYFDNFIIC SEKEVADHWA ADGWRWKIMI ANANKPGVLK QLMAAAEGHP WLWLIYLVTA 480
       GVPIALITSF CWPRKVKKKH KDTEYKKTDI CIPQTKGVLE QEEKEEKAAL EKPMDLEEEK 540
       KQNDGEMLEK EEESEPEEKS EEEIEIIEGQ EESNQSNKSG SEDEMKEADE STGSGDGPIK 600
       SVRKRRVRKD

                                        SEQ ID NO:161 PEZ9 DNA SEQUENCE
25   Nucleic Acid Accession #:   NM_005932
       Coding sequence: 75-2216 (underlined sequences correspond to start and stop codons)

30   1   11   21   31   41   51
       | | | | | |
       GCGGAGCGCG CGCTCCCAGC GAAAGCAGCA GGGCAGGGAT CTGCGTTGGA GGAAGGGACT   60
       GCTCTGGTGC TAGAATGCTG TGCGTCGGAA GGCTGGGCGG CTTGGGAGCC AGAGCAGCAG 120
       CTCTGCCGCC CCGCCGGGCG GGCCGGGGAA GCCTCGAAGC CGGGATCCGG GCCCGAAGGG 180
       TCAGCACCAG CTGGTCTCCC GTGGGCGCCG CCTTCAATGT CAAGCCCCAG GGCAGCCGCT 240
35   TGGACCTGTT CGGCGAGCGG GCGCGTCTTT TTGGAGTTCC TGAGCTGAGT GCCCCAAAG 300
       GATTTCATAT TGCACAAGAA AAAGCCTTGA GAAAGACAGA ATTGCTTGTG GACCGTGCAT 360
       GTTCCACCCC ACCTGGGCCC CAGACCGTGC TGATCTTCGA TGAGCTCTCG GATTCCTTAT 420
       GCAGAGTGGC CGACTTGGCT GATTTTGTGA AAATCGCTCA CCCTGAGCCA GCATTCAGAG 480
       AAGCTGCGGA AGAAGCTTGT AGAAGTATTG GCACCATGGT AGAGAAGTTG AACACAAATG 540
40   TGGATTTATA TCAAAGTTTG CAAAAATTAC TAGCTGATAA AAAACTTGTG GATTCCCTTG 600
       ATCCAGAAAC AAGGCGAGTG GCTGAACTGT TTATGTTTGA TTTTGAAATT AGTGGAATCC 660
       ATCTAGACAA ACAAAAGCGT AAAAGAGCAG TGGACCTCAA TGTTAAAATC TTGGATTTGA 720
       GTAGTACATT TCTTATGGGA ACCAATTTTC CCAACAAGAT TGAGAAGCAT CTCTTACCAG 780
       AACACATTCG TCGTAACTTT ACATCTGCTG GGGATCATAT CATAATTGAT GGTCTCCACG 840
45   CAGAATCACC AGATGACTTG GTGCGAGAAG CTGCTTATAA AATTTTTCTT TATCCCAATG 900
       CTGGTCAATT GAAATGTTTA GAAGAATTGC TCAGCAGCAG AGATCTTCTG GCAAAGTTGG 960
       TGGGGTATTC CACGTTTTCT CACAGGGCTC TCCAAGGAAC GATAGCTAAA AATCCAGAGA 1020
       CTGTCATGCA GTTCCTTGAA AAACTATCTG ACAAACTTTC TGAAAGAACT CTGAAAGATT 1080
       TTGAGATGAT ACGAGGGATG AAAATGAAAC TGAATGCTCA AAATTCCGAA GTAATGCCCT 1140
       GGGACCCCCC TTACTACAGT GGTGTGATTC GTGCAGAAAG GTATAATATT GAGCCCAGCC 1200
50   TATATTGCCC GTTTTTCTCT CTTGGAGCAT GCATGGAAGG CCTGAATATT TTGCTTAACA 1260
       GACTGTTGGG GATTTCATTA TATGCAGAGC AGCCTGCAAA GGAGGAGGTG TGGAGCGAAG 1320
       ATGTCCGAAA ACTGGCTGTT GTTCATGAAT CTGAAGGATT GTTGGGGTAC ATTTACTGTG 1380
       ATTTTTTTCA GCGAGCAGAC AAACCACATC AGGATTGCCA TTTCACTATC CGTGGAGGCA 1440
55   GACTAAAGGA AGATGGAGAC TATCAACTCC CACTTGTAGT TCTTATGCTG AATCTTCCCC 1500
       GTTCCTCAAG GAGTTCTCCA ACTTTGCTAA CTCCTGGCAT GATGGAAAAT CTTTTCCATG 1560
       AAATGGGACA TGCCATGCAT TCAATGCTAG GACGTACTCG TTACCAACAC GTCACTGGGA 1620
       CCAGGTGCCC TACTGATTTT GCTGAGGTTC CTTCTATTCT GATGGAGTAC TTTGCAAATG 1680
       ATTATCGAGT AGTTAACCAA TTTGCCAGAC ATTATCAGAC TGGACAGCCA CTGCCAAAAA 1740
60   ATATGGTGTC TCGTCTTTGT GAATCTAAAA AGGTTTGTGC TGCAGCTGAT ATGCAACTTC 1800
       AGGTCTTTTA TGCCACTCTG GATCAAATCT ACCATGGGAA GCATCCCCTG AGGAATTCAA 1860
       CCACAGACAT TCTCAAGGAA ACACAAGAGA AATTCTATGG CCTACCATAT GTTCCAAATA 1920
       CTGCCTGGCA GCTGCGATTC AGCCACCTCG TGGGGTATGG TGCTAGATAT TACTCTTACC 1980
       TCATGTCCAG AGCGGTCGCC TCCATGGTTT GGAAGGAGTG TTTTCTACAG GATCCTTTCA 2040
65   ACAGGGCTGC CGGGGAGCGC TATCGCAGGG AGATGCTGGC CCACGGTGGA GGCAGGGAGC 2100
       CCATGCTCAT GGTTGAAGGT ATGCTTCAGA AGTGTCCTTC TGTTGATGAC TTCGTAAGTG 2160
       CCCTCGTTTC CGACTTGGAT CTGGACTTCG AAACTTTCCT CATGGATTCT GAATAAAAGA 2220
       AACACTCTAC ACCTCTAATC AAGGTCATGT AGTAATGACT TTGTTATAAA TGCTACAGCT 2280
       GTGAGAGCTT GTTTCTGATT GTTTCATTGT TCGCTTCTGT AATTCTGAAA AACTTTAAAC 2340
70   TGGTAGAACT TGGAATAAAT AATTTGTTTT AATTAAAAAA AAAAAAAAAA AA

SEQ ID NO:162 PEZ9 Protein sequence:
Protein Accession #:   NP_005923.1

75   1   11   21   31   41   51
       | | | | | |
       MLCVGRLGGL GARAAALPPR RAGRGSLEAG IRARRVSTSW SPVGAAFNVK PQGSRLDLFG   60
       ERARLFGVPE LSAPEGFHIA QEKALRKTEL LVDRACSTPP GPQTVLIFDE LSDSLCRVAD 120

                                       364

LADFVKIAHP EPAFREAAEE ACRSIGTMVE KLNTNVDLYQ SLQKLLADKK LVDSLDPETR 180
RVAELFMFDF EISGIHLDKQ KRKRAVDLNV KILDLSSTFL MGTNFPNKIE KHLLPEHIRR 240
NFTSAGDHII IDGLHAESPD DLVREAAYKI FLYPNAGQLK CLEELLSSRD LLAKLVGYST 300
FSHRALQGTI AKNPETVMQF LEKLSDKLSE RTLKDFEMIR GMKMKLNAQN SEVMPWDPPY 360
YSGVIRAERY NIEPSLYCPF FSLGACMEGL NILLNRLLGI SLYAEQPAKG EVVWSEDVRKL 420
AVVHESEGLL GYIYCDFFQR ADKPHQDCHF TIRGGRLKED GDYQLPLVVL MLNLPRSSRS 480
SPTLLTPGMM ENLFHEMGHA MHSMLGRTRY QHVTGTRCPT DFAEVPSILM EYFANDYRVV 540
NQFARHYQTG QPLPKNMVSR LCESKKVCAA ADMQLQVFYA TLDQIYHGKH PLRNSTTDIL 600
KETQEKFYGL PYVPNTAWQL RFSHLVGYGA RYYSYLMSRA VASMVWKECF LQDPFNRAAG 660
ERYRREMLAH GGGREPMLMV EGMLQKCPSV DDFVSALVSD LDLDFETFLM DSE

SEQ ID NO:163 PEZ8 DNA SEQUENCE

Nucleic Acid Accession #:   AF103907
Coding sequence: none (underlined sequences correspond to start and stop codons)

```
1     11    21    31    41    51
|     |     |     |     |     |
ACAGAAGAAA TAGCAAGTGC CGAGAAGCTG GCATCAGAAA AACAGAGGGG AGATTTGTGT   60
GGCTGCAGCC GAGGGAGACC AGGAAGATCT GCATGGTGGG AAGGACCTGA TGATACAGAG   120
GAATTACAAC ACATATACTT AGTGTTTCAA TGAACACCAA GATAAATAAG TGAAGAGCTA   180
GTCCGCTGTG AGTCTCCTCA GTGACACAGG GCTGGATCAC CATCGACGGC ACTTTCTGAG   240
TACTCAGTGC AGCAAAGAAA GACTACAGAC ATCTCAATGG CAGGGGTGAG AAATAAGAAA   300
GGCTGCTGAC TTTACCATCT GAGGCCACAC ATCTGCTGAA ATGGAGATAA TTAACATCAC   360
TAGAAACAGC AAGATGACAA TATAATGTCT AAGTAGTGAC ATGTTTTTGC ACATTTCCAG   420
CCCCTTTAAA TATCCACACA CACAGGAAGC ACAAAAGGAA GCACAGAGAT CCCTGGGAGA   480
AATGCCCGGC CGCCATCTTG GGTCATCGAT GAGCCTCGCC CTGTGCCTGG TCCCGCTTGT   540
GAGGGAAGGA CATTAGAAAA TGAATTGATG TGTTCCTTAA AGGATGGGCA GGAAAACAGA   600
TCCTGTTGTG GATATTTATT TGAACGGGAT TACAGATTTG AAATGAAGTC ACAAAGTGAG   660
CATTACCAAT GAGAGGAAAA CAGACGAGAA AATCTTGATG GCTTCACAAG ACATGCAACA   720
AACAAAATGG AATACTGTGA TGACATGAGG CAGCCAAGCT GGGGAGGAGA TAACCACGGG   780
GCAGAGGGTC AGGATTCTGG CCCTGCTGCC TAAACTGTGC GTTCATAACC AAATCATTTC   840
ATATTTCTAA CCCTCAAAAC AAAGCTGTTG TAATATCTGA TCTCTACGGT TCCTTCTGGG   900
CCCAACATTC TCCATATATC CAGCCACACT CATTTTTAAT ATTTAGTTCC CAGATCTGTA   960
CTGTGACCTT TCTACACTGT AGAATAACAT TACTCATTTT GTTCAAAGAC CCTTCGTGTT   1020
GCTGCCTAAT ATGTAGCTGA CTGTTTTTCC TAAGGAGTGT TCTGGCCCAG GGATCTGTG   1080
AACAGGCTGG GAAGCATCTC AAGATCTTTC CAGGGTTATA CTTACTAGCA CACAGCATGA   1140
TCATTACGGA GTGAATTATC TAATCAACAT CATCCTCAGT GTCTTTGCCC ATACTGAAAT   1200
TCATTTCCCA CTTTTGTGCC CATTCTCAAG ACCTCAAAAT GTCATTCCAT TAATATCACA   1260
GGATTAACTT TTTTTTTTAA CCTGGAAGAA TTCAATGTTA CATGCAGCTA TGGGAATTTA   1320
ATTACATATT TTGTTTTCCA GTGCAAAGAT GACTAAGTCC TTTATCCCTC CCCTTTGTTT   1380
GATTTTTTTT CCAGTATAAA GTTAAAATGC TTAGCCTTGT ACTGAGGCTG TATACAGCAC   1440
AGCCTCTCCC CATCCCTCCA GCCTTATCTG TCATCACCAT CAACCCCTCC CATACCACCT   1500
AAACAAAATC TAACTTGTAA TTCCTTGAAC ATGTCAGGAC ATACATTATT CCTTCTGCCT   1560
GAGAAGCTCT TCCTTGTCTC TTAAATCTAG AATGATGTAA AGTTTTGAAT AAGTTGACTA   1620
TCTTACTTCA TGCAAAGAAG GGACACATAT GAGATTCATC ATCACATGAG ACAGCAAATA   1680
CTAAAAGTGT AATTTGATTA TAAGAGTTTA GATAAATATA TGAAATGCAA GAGCCACAGA   1740
GGGAATGTTT ATGGGGCACG TTTGTAAGCC TGGGATGTGA AGCAAAGGCA GGGAACCTCA   1800
TAGTATCTTA TATAATATAC TTCATTTCTC TATCTCTATC ACAATATCCA ACAAGCTTTT   1860
CACAGAATTC ATGCAGTGCA AATCCCCAAA GGTAACCTTT ATCCATTTCA TGGTGAGTGC   1920
GCTTTAGAAT TTTGGCAAAT CATACTGGTC ACTTATCTCA ACTTTGAGAT GTGTTTGTCC   1980
TTGTAGTTAA TTGAAAGAAA TAGGGCACTC TTGTGAGCCA CTTTAGGGTT CACTCCTGGC   2040
AATAAAGAAT TTACAAAGAG CTACTCAGGA CCAGTTGTTA AGAGCTCTGT GTGTGTGTGT   2100
GTGTGTGTGT GAGTGTACAT GCCAAAGTGT GCCTCTCTCT CTTGACCCAT TATTTCGACG   2160
TTAAAACAAG CATGTTTTCA AATGGCACTA TGAGCTGCCA ATGATGTATC ACCACCATAT   2220
CTCATTATTC TCCAGTAAAT GTGATAATAA TGTCATCTGT TAACATAAAA AAAGTTTGAC   2280
TTCACAAAAG CAGCTGGAAA TGGACAACCA CAATATGCAT AAATCTAACT CCTACCATCA   2340
GCTACACACT GCTTGACATA TATTGTTAGA AGCACCTCGC ATTTGTGGGT TCTCTTAAGC   2400
AAAATACTTG CATTAGGTCT CAGCTGGGGC TGTGCATCAG GCGGTTTGAG AAATATTCAA   2460
TTCTCAGCAG AAGCCAGAAT TTGAATTCCC TCATCTTTTA GGAATCATTT ACCAGGTTTG   2520
GAGAGGATTC AGACAGCTCA GGTGCTTTCA CTAATGTCTC TGAACTTCTG TCCCTCTTTG   2580
TGTTCATGGA TAGTCCAATA AATAATGTTA TCTTTGAACT GATGCTCATA GGAGAGAATA   2640
TAAGAACTCT GAGTGATATC AACATTAGGG ATTCAAAGAA ATATTAGATT TAAGCTCACA   2700
CTGGTCAAAA GGAACCAAGA TACAAAGAAC TCTGAGCTGT CATCGTCCCC ATCTCTGTGA   2760
GCCACAACCA ACAGCAGGAC CCAACGCATG TCTGAGATCC TTAAATCAAG GAAACCAGTG   2820
TCATGAGTTG AATTCTCCTA TTATGGATGC TAGCTTCTGG CCATCTCTGG CTCTCCTCTT   2880
GACACATATT AGCTTCTAGC CTTTGCTTCC ACGACTTTTA TCTTTTCTCC AACACATCGC   2940
TTACCAATCC TCTCTCTGCT CTGTTGCTTT GGACTTCCCC ACAAGAATTT CAACGACTCT   3000
CAAGTCTTTT CTTCCATCCC CACCACTAAC CTGAATGCCT AGACCCTTAT TTTTATTAAT   3060
TTCCAATAGA TGCTGCCTAT GGGCTATATT GCTTTAGATG AACATTAGAT ATTTAAAGCT   3120
CAAGAGGTTC AAAATCCAAC TCATTATCTT CTCTTTCTTT CACCTCCCTG CTCCTCTCCC   3180
TATATTACTG ATTGCACTGA ACAGCATGGT CCCCAATGTA GCCATGCAAA TGAGAAACCC   3240
AGTGGCTCCT TGTGGTACAT GCATGCAAGA CTGCTGAAGC CAGAAGGATG ACTGATTACG   3300
CCTCATGGGT GGAGGGGACC ACTCCTGGGC CTTCGTGATT GTCAGGAGCA AGACCTGAGA   3360
TGCTCCCTGC CTTCAGTGTC CTCTGCATCT CCCCTTTCTA ATGAAGATCC ATAGAATTTG   3420
CTACATTTGA GAATTCCAAT TAGGAACTCA CATGTTTTAT CTGCCCTATC AATTTTTTAA   3480
ACTTGCTGAA AATTAAGTTT TTTCAAAATC TGTCCTTGTA AATTACTTTT TCTTACAGTG   3540
TCTTGGCATA CTATATCAAC TTTGATTCTT TGTTACAACT TTTCTTACTC TTTTATCACC   3600
```

365

AAAGTGGCTT TTATTCTCTT TATTATTATT ATTTTCTTTT ACTACTATAT TACGTTGTTA 3660
TTATTTTGTT CTCTATAGTA TCAATTTATT TGATTTAGTT TCAATTTATT TTTATTGCTG 3720
ACTTTTAAAA TAAGTGATTC GGGGGGTGGG AGAACAGGGG AGGGAGAGCA TTAGGACAAA 3780
TACCTAATGC ATGTGGGACT TAAAACCTAG ATGATGGGTT GATAGGTGCA GCAAACCACT 3840
ATGGCACACG TATACCTGTG TAACAAACCT ACACATTCTG CACATGTATC CCAGAACGTA 3900
AAGTAAAATT TAAAAAAAAG TGA

PEZ8 Protein sequence:
Protein Accession #:   none

SEQ ID NO:164 PEZ6 DNA SEQUENCE

Nucleic Acid Accession #:   AB028945
Coding sequence: 1-3765 (underlined sequences correspond to start and stop codons)

```
  1     11    21    31    41    51
  |     |     |     |     |     |
```

ATGATGATGA ACGTCCCCGG CGGAGGAGCG GCCGCGGTGA TGATGACGGG CTACAATAAT 60
GGTCGCTGTC CCCGGAATTC TCTCTACAGT GACTGCATTA TTGAGGAGAA GACGGTGGTC 120
CTGCAGAAAA AAGACAATGA GGGCTTTGGA TTCGTGCTTC GAGGGGCCAA AGCTGACACA 180
CCCATTGAAG AATTCACACC AACACCGGCT TTCCCAGCCC TACAGTACCT GGAGTCCGTG 240
GATGAAGGTG GGGTGGCGTG GCAAGCCGGA CTAAGGACCG GGGACTTCTT GATTGAGGTT 300
AACAATGAGA ATGTTGTCAA AGTCGGCCAC AGGCAGGTGG TGAACATGAT CCGGCAGGGA 360
GGGAATCACC TGGTCCTTAA GGTGGTCACG GTGACCAGGA ATCTGGACCC CGACGACACC 420
GCCAGGAAGA AAGCTCCCCC GCCTCCAAAG CGGGCACCGA CCACAGCCCT CACCCTGCGC 480
TCCAAGTCCA TGACCTCGGA GCTGGAGGAG CTCGTGGATA AAGATAAACC CGAGGAGATA 540
GTCCCGGCCT CCAAGCCCTC CCGCGCTGCT GAGAACATGG CTGTGGAACC GAGGGTGGCG 600
ACCATCAAGC AGCGGCCCAG CAGCCGGTGC TTCCCGGCGG GCTCAGACAT GAACTCTGTG 660
TACGAACGCC AAGGAATCGC CGTGATGACG CCCACTGTTC CTGGGAGCCC AAAAGCCCCG 720
TTTCTGGGCA TCCCTCGAGG TACGATGCGA AGGCAGAAAT CAATAGACAG CAGAATCTTT 780
CTATCAGGAA TAACAGAGGA AGAGCGGCAG TTTCTGGCTC CTCCAATGCT GAAGTTCACC 840
AGAAGCCTGT CCATGCCGGA CACCTCTGAG GACATCCCCG CTCCACCGCA GTCTGTGCCC 900
CCGTCCCCAC CACCACCTTC CCCAACCACT TACAACTGCC CCAAGTCCCC AACTCCAAGA 960
GTCTACGGGA CGATTAAGCC TGCCGTTCAAT CAGAATTCTG CCGCCAAGGT GTCCCCCGCC 1020
ACCAGGTCCG ACACCGTGGC CACCATGATG AGGGACGAAGG GGATGTACTT CAGGAGAGAG 1080
CTGGACCGCT ACTCCTTGGA CTCTGAAGAC CTCTACAGTC GGAATGCCGG CCCGCAAGCC 1140
AACTTCCGCA ACAAGAGAGG CCAGATGCCA GAAAACCCAT ACTCAGAGGT GGGGAAGATC 1200
GCCAGCAAAG CCGTCTACGT CCCCCGCCAAG CCCGCCAGGC GGAAGGGGAT GCTGGTGAAG 1260
CAGTCCAACG TGGAGGACAG CCCCGAGAAG ACGTGCTACA TCCCTATCCC GACCATCATC 1320
GTGAAGGAGC CGTCCACCAG CAGCAGCGGC AAGAGCAGCC AGGGCAGCAG CATGGAGATC 1380
GACCCCCAGG CCCCGGAGCC ACCGAGCCAG CTGCGGCCTG ACGAAAGCCT GACCGTCAGC 1440
AGCCCCTTTG CCGCCGCCAT CGCCGGAGCC GTCCGCGACC GTGAGAAGCG GCTGGAAGCC 1500
AGGAGGAACT CCCCGGCCTT CCTCTCCACA GACCTGGGGG ATGAGGATGT GGGCCTGGGG 1560
CCACCCGCCC CCAGGACGCG GCCCTCCATG TTCCCCGAGG AGGGGGGATTT TGCTGACGAG 1620
GACAGCGCTG AGCAGCTGTC ATCCCCCATG CCGAGTGCCA CGCCCAGGGA GCCCGAAAAC 1680
CATTTCGTGG GTGGCGCCGA GGCCAGTGCT CCGGGTGAGG CTGGGAGGCC GCTGAATTCC 1740
ACGTCCAAAG CCCAGGGGCC CGAGAGCAGC CCAGCAGTGC CCTCCGCGAG CAGCGGCACA 1800
GCCGGCCCCG GGAATTATGT CCACCCACTC ACAGGGCGGC TGCTTGATCC CAGCTCCCCG 1860
CTGGCCCTGG CACTCTCCGC AAGGGACCGA GCCATGAAGG AGTCTCAACA GGGACCCAAA 1920
GGGGAGGCCC CCAAGGCCGA CCTCAACAAA CCTCTTTACA TTGATACCAA AATGCGGCCC 1980
AGCCTGGATG CCGGCTTCCC TACGGTCACC AGGCAGAACA CCCGGGGGACC CCTGAGGCGG 2040
CAGGAGACGG AGAACAAGTA CGAGACCGAC CTGGGCCGAG ACCGGAAAGG CGATGACAAG 2100
AAGAACATGC TGATCGACAT CATGGACACG TCCCAGCAAG AGTCGGCTGG CCTGCTGATG 2160
GTGCACACCG TGGACGCCAC TAAGCTGGAC AACGCCCTGC AGGAAGAGGA CGAGAAGGCA 2220
GAGGTGGAGA TGAAGCCAGA CAGCTCGCCG TCCGAGGTGC CAGAAGGTGT TTCCGAAACC 2280
GAAGGTGCTT TACAGATCTC CGCTGCCCCC GAGCCCACCA CCGTGCCCGG CAGAACCATC 2340
GTCGCGGTGG GCTCCATGGA AGAGGCGGTG ATTTTGCCAT TCCGCATCCC TCCTCCCCCT 2400
CTGGCATCCG TGGACTTGGA TGAGGATTTT ATTTTTACAG AGCCATTGCC TCCTCCCCTG 2460
GAATTTGCAA ATAGTTTTGA TATCCCCGAT GACCGGGCAG CTTCTGTCCC GGCTCTCTCA 2520
GACTTAGTGA AGCAGAAGAA AAGCGACCAC CCTCAGTCCC CTTCGTTGAA CTCCAGCCAA 2580
CCAACCAACT CTGCAGACAG CAAGAAGCCA GCCAGTCTTT CAAACTGTCT GCCTGCCTCA 2640
TTCCTGCCAC CCCCTGAAAG CTTTGACGCC GTCGCCGACT CTGGGATCGA GGAGGTGGAC 2700
AGCCGGAGTA GCAGCGACCA CCACCTCGAG ACGACCAGCA CTATCTCCAC CGTGTCTAGC 2760
ATCTCCACCC TGTCTTCCGA AGGTGGAGAG AATGTGGACA CCTGCACAGT CTATGCAGAT 2820
GGGCAAGCAT TTATGGTTGA CAAACCCCCA GTACCTCCTA AGCCAAAAAT GAAGCCCATC 2880
ATTCACAAAA GCAATGCACT TTATCAAGAC GCGCTCGTGG AAGAAGATGT AGATAGCTTT 2940
GTTATCCCCC CGCCCCGCTCC CCCGACCCCCG CCGGGCAGTG CCCAGCCTGG GATGGCCAAG 3000
GTTCTCCAGC CAAGGACCTC CAAGTTGTGG GGCGACGTCA CAGAGATCAA AAGCCCGATT 3060
CTCTCAGGCC CAAAGGCAAA CGTTATTAGT GAATTGAACT CTATCCTACA GCAAATGAAC 3120
CGAGAGAAAT TGGCAAAGCC GGGGGAAGGA CTGGATTCAC CAATGGGAGC CAAGTCCGCC 3180
AGCCTCGCTC CAAGAAGCCC GGAGATCATG AGCACCATCT CAGGTACACG GAGCACGACG 3240
GTCACCTTCA CTGTTCGCCC CGGCACCTCC CAGCCCATCA CCCTGCAGAG CCGGCCCCCC 3300
GACTATGAAA GCAGGACCTC AGGAACAAGA CGTGCCCCAA GCCCTGTGGT CTCGCCAACA 3360
GAGATGAACA AAGAGACCCT GCCCGCCCCC CTGTCTGCTG CCACCGACCTC TCCTTCTCCC 3420
GCTCTCTCAG ATGTCTTTAG CCTTCCAAGC CAGCCCCCTT CTGGGGATCT ATTTGGCTTG 3480
AACCCAGCGG GACGCAGTAG GTCGCCATCC CCCTCGATAC TGCAACAGCC AATCTCAAAT 3540
AAGCCTTTTA CAACTAAACC TGTCCACCTG TGGACTAAAC CAGATGTGGC CGATTGGCTG 3600
GAAAGTCTAA ACTTGGGTGA ACATAAAGAG GCCTTCATGG ACAATGAGAT CGATGGCAGT 3660
CACTTACCAA ACCTGCAGAA GGAGGACCTC ATCGATCTTG GGGTAACTCG AGTCGGGCAC 3720

366

```
     AGAATGAACA TAGAAAGGGC TTTGAAACAG CTGCTGGACA GATAAGGACG GCTGCTCTCC 3780
     ACCTCGCAGA CTGCTCTTGT TATAAGTAGA GATGGGCTCG TGCTGAAACA TCTGAATGCC 3840
     AAGCGAAGTC TGTGAGCATC AACCCCACTC CATGGGTTTG TCTCCTGGTA CCCAAAGAAA 3900
5    TACTGAGTTG TGTCCACAAC ATGGCTGGGT CTTCAGACCC CTGGCTCACC ATGTGGGTGT 3960
     CTTGGGCAGT TTCTATCACA CATGGGACAA GGGGAGGGAG TTTTTCTAAC ATGGAAAAAG 4020
     ATTCCCAGCC TGCCGCCCAG CATGCAGGTG GCCTCGCTTT GCCGGGTCCG AGAGGCTCCC 4080
     CGTCAATTTT GCACGGGATC CTAGCTCTTG TAGGCAGACA CCAGTGCACT CTAGATACCT 4140
     CCTGAGACCT CCGTCCTCTG CTTTCCGGGC AGCTCTCACC ACCCCAGGCC CCGGCATGAG 4200
     GCCTTTCCTC AGTCCTGTGG CCTCTCAGAG GACACCTGAT GCTCACCTGC CCCTCTTTCT 4260
10   CCTGCACTTG GCTTGCAGTG AGATGCTCCC AGATGCATTT GTCCAGTGCC CCATCATGGG 4320
     CCTGAAAGGC AGAGAAACTT TTTCCTACAC AGATTCTTTT CCCCATCTCC TCCTGTGGTT 4380
     TGCATCCATG GCTCTTTGGC CATGAGGTTC CTGGCAGTGC TGGGAGTTTG GATGGGATCG 4440
     TGCCCAGCTT TGCTTAGCTT TCTTTATTTC TGCAAATCTG TTAGCATAAT TCCAAGGTGG 4500
15   CCAAGCAGAT GTCACATGGA GTTAGTCAAA GCACAAAGTC ACGATTCCAC AATGGAGGGG 4560
     AGACCTGGCC AAGGGAGCCA GCCAGCGTGC AACTGCCCAA GCTCCAGGTC TCCAGGACAA 4620
     GAGCAGTTGT CTGCCATGAG CACCCATCCA GGATGGAGAA TAAGGGCTTC TCTGCCTCTC 4680
     AGAATTCTTT TTAATTGAAG ATGTCTTGAG CTCTGCAAAG ATCAGACAG GTGAGCATCC 4740
     ACTTTGACAT GAAGGACAAG AAGACGCATG GCTCATGGCG GGCACATGCG GCTGCCAGTG 4800
20   AGACAGCGTC TCCTCTGGGA GCTGGGCGGG CACAGCATCC TCAGTTCTGT GCCCAGCCAA 4860
     GGGTGAGCAT CTCTGCTGAG ACAGTCCTTT TGCTCTCGGA GGCCAGGGAA GATGGTACTT 4920
     AGAGGCTTTT CCCCTATCGC TCTGGGTGTC TAGGAATCCC ACCAGCTTGT CTTAACAGTA 4980
     CAACAGCTTC TTTGAGGACC CAGTGGGTAT GGAGTATAGA CAGAACCCAG GGGTTGAGAAC 5040
     AGAAGGTGGG CGGCAGGATC AGAGTGAAAG CAGAGGCGTG AGGAGAGGAA AGCAGGGAGG 5100
     TCTCCTGGGC TGCCAGGTCA GCCTCTCTGG CAAGGCTTTC TTGAGCCCCG CCCCTTTCTT 5160
25   TCCCCGGAGT CCCTCCACCC CATAACAATA CCTCGAATTT CCAAAAGAGG TCACCAGATG 5220
     CACATGGGCC GCAAAACACA CAGTCAGGCT TCCAGCACAT TCTCCCCCAT TTGGAGGATA 5280
     CTCGAATGTC AGGTTTTTGG TTTTATTATT ATTTCAGAAC TAGCTCAGCC CATCTCTAAT 5340
     TATAAAACAT GGTTTTGTTT TTTTTTTTTC CTTTTTTTCT TGATTAGGTC TGGAACAGCT 5400
30   CTAGAATGAA CACATAAAAT TTAGCAATTT AAAATCTTTC TTTACTGCAA GTTTAAATAG 5460
     TTGTACAGAT AGTTTATAAG CACAATATTT TAAGAAAAAA AAGTGGCTGG TCTACTAGGC 5520
     AGCCTTTGTG CCACTTCAGT GCTAGAAAGT TAAAGAAAAA AAAACTTTTG TGATTTAATA 5580
     ATACTATTTC TGTGGAATAA TTATAAAAGT ATGACCTTTT TAAATCAACC TTATTTGGAT 5640
     GCATCTGAAC CAGCAGAGCT GTGTTATATT TTCTATCTTT GCTAGAACTT CGTCATTGAA 5700
35   GGACAATTTC TTCAAAGTGG TTACAATTCA TAATGCAGCA GTTTCTCCAA AAACAAAAAC 5760
     AAAACACACA CCACACACAC GCGCTTTTCC AGTCACACAC CCCTGATGTT GGAACCAAGT 5820
     TTTTGGACCT TCTGTTCCAA AACCTTTTGC AGGTCAATCT TTGTATTTGA AATGATCCAA 5880
     TCCAACTTGA AGTCAATTGA ATATTAAGGC GCTTTACTTC CGTGTGCTTT CAGTTTTTCC 5940
     ATCATGAGAT GAATGAGCAT TACTCTAGAT AAATTTCAAG ACAGGATACT ACAGGTGGCC 6000
40   TGCTGAGGCT GCCCCATATT TTAGAAAATG TAAAAATGGT GGTTTGGCCA TTAATTTGTC 6060
     TTCCATTTGA TGATACCGCA AAATTCCGTG AGTCCATTCC TTTGGCATGG CACTTTCCCT 6120
     GGGCCTACAG TTGGTATTAC CTCTGTGCTC AGTGCCAGGC AAAACACTAG CTCAAAGGAG 6180
     AGTCAAGGAA ACCGCTGGCA GACGATAACC AGTCGAAACT CGTGACTTCG GTTTGTTGAA 6240
     CTTTGGCAGC CAGTTGGTGA GGGCCAGATG TTATTCCCTT TCTTAAAGAT ACTCCAAGCC 6300
45   ACATGCCACT AACCACAAGC AAGCTGGCTG CAAGACTAAA GAGCTGATAA CATAGTTTAT 6360
     TTTTACACTG TCTTATTATA GAGAAGTAAT AGACCTATCA GAACCTGCAC TGACCAACAA 6420
     ATAAACACAT GTTGCCAAGA TGAATCGGTC TCTATCTCTA TCTGCTTATT TTGGTACTGA 6480
     AAGCAATAGT TCCTCATTCA AATCACCACC CACTGTTCTC CCCCTTTGGG ACATGTTAGG 6540
     ACGAGGCCCT ATTCCATGCC CCTCTTTAAT GGTGGAACAA ATGTTAAACT GCTCATCTAA 6600
     AGATCATGTT GATATTATTC CAGGTTTTAA GATCAACTTT TGTTACATAC TGTAATTTAA 6660
50   ATAAACTGCA TTTACATGCC TAGTTTCTGT AATATTGTGT ATACAAAACC CAAATCTCTC 6720
     AAAATGTAAA TTATGTATAC CTGCCAAGAT ACCTTTTCCA GGGTGTCTGC GCACATTTTA 6780
     AGTTAATTCA CATAATATAA AAATTACTCA ATGTGACTGT TGATTTGCTG AACTTTACAT 6840
     ATCACAAAGT GAATTATTTG TGATACTTTA GTTAATAAAA TGGTAAATTT TTTTCTCAGT 6900
     TATTGAACAA GCAAGCATTA TCCAGTTGAT CTGGCAATGA CTTTTTGTGT GTGGGCCACA 6960
55   ATATTGATTT TCCCATTAAC AATTTTTTTT TGTTTTTTAA ATACTAATAT GTTTCACACT 7020
     ATAGTTTGTG TAACAACACG TGTTCGCATT ATCTATGTTG CTGTTACTTT TGTGCTTTTA 7080
     TTCTTTTTAG ACTTTATAAA AAAAAAAAAA AGCTCCTGTA ATTTGCACTT TCTCCCAATC 7140
     CTTAAATCTC TTGTATGGCA ACCAAAATTA CTGTAAAAAA ATAAATATAC TATTGCACTA 7200
     .AGGTTGTGGT TCTGATTGCA AACAAACAGT GAACACTGTC TGAATTAAAC AAAAAGCTGC 7260
60   CCGACTTGCA ATCTAATGTA GATTATCTCA GGCATTGTGG CCAGCTCTGC CTCTCTAAAA 7320
     CTGACCAGAA AAATCTCTCT CATCGAGTAA ACAGGCTCCT GTCACTGAGC TAATCTGCCT 7380
     TGGTTCCATT TCCTTATTCT CAATTTATCA ATGGATACGT GCATGTTATT TCAGAATTAT 7440
     GCAAAACGTC AAAATCTGCT TCTGTGACCG CTGCTATAGG CGTGGAGCTG AGGCTCGGCT 7500
     TTTCCTTTTG TTCTGGGTGG AAGCAGCGGT GCCGCGGAGG GCCAGCCAGA TCCGGACCCT 7560
65   TCCCTTAGGG TCCAGTCTCC CCACACCCCA GCAGGGTGTC TTCTAGCCAT AAGGCCAAGG 7620
     GAGTGGCAGA ACTGGGCCGC CTCTCTGGTT GACAAGCAAA CCACATGCTA AGGCTTGGAG 7680
     CAAGAGAGAA TTTGTGTCTA TTGGCAAAGA ACTAAGCCAG GAAGACATGG GCCATCCCTC 7740
     CGCTTTAGGG AAGCATATTT TAAACCTAAA CGTTGAACTT CTTCTTTGGC CTCACCAGTG 7800
     AAAACTTGTT GTCTTTAGTT CCTAAAGTTT CTTCTACTTT GGCACATTCC CCAGTTGAGC 7860
70   AGCAGCCTCT ATGCTTCCAC GTTCAGGAAA AATTCCAGTC CTCATATCTT TTGTAGTTCA 7920
     CCCTCAAGCT CTCCCGCTTC ACCATCCAAT AGTTTCTCCC AAACCTTGGC ACCCCCCTAG 7980
     ACTTTGCTTC CAATGGTTTC TTCCAGACCA CTTTTCCTAG ATGAATATAT TCGTTTACCT 8040
     TACTAGGAAA ATTATTGGAA GATTTTTTCT TTTACTTGAA ATTGGAGGCA TTTTAATAAC 8100
     TGGCGAACTG GAATGTGTTT CTGTATTTGT AGACAACCAT GTACCCATGC AAGTAGGTGA 8160
75   ACATTCCACA GTGGCTGGGT GACCACAGCA GCTGCATGCA GACAGGACTG CCCGTGCTTT 8220
     GTGGGGAATC AGAGAATTTC CAAACTTGTT TCTCAGACTT CCGCAGATCT CATCACTTTG 8280
     ATTTCTAATC CATGCTGTAT TGGTGATTTT GTTTATCGTT CCTGTAACTT GTTCTACATT 8340
     CCACAGTCTT TACCGTTTTA TGTTCAAAAT TACAACAATC CCTGTCCATT GATTCCACTC 8400
     TGGAACTCTT TGTTCATGCC AATTTTGAAA TTTTAATACG AGCCTTCAAA TAAACACAGA 8460
```

367

AAAGAAAAAA AAAAAAAAAA AAAAAAAA

<u>SEQ ID NO:165 PEZ6 Protein sequence:</u>
Protein Accession #:   BAA82974.1

```
       1     11    21    31    41    51
       |     |     |     |     |     |
MMMNVPGGGA AAVMMTGYNN GRCPRNSLYS DCIIEEKTVV LQKKDNEGFG FVLRGAKADT  60
PIEEFTPTPA FPALQYLESV DEGGVAWQAG LRTGDFLIEV NNENVVKVGH RQVVNMIRQG  120
GNHLVLKVVT VTRNLDPDDT ARKKAPPPPK RAPTTALTLR SKSMTSELEE LVDKDKPEEI  180
VPASKPSRAA ENMAVEPRVA TIKQRPSSRC FPAGSDMNSV YERQGIAVMT PTVPGSPKAP  240
FLGIPRGTMR RQKSIDSRIF LSGITEEERQ FLAPPMLKFT RSLSMPDTSE DIPPPPQSVP  300
PSPPPPSPTT YNCPKSPTPR VYGTIKPAFN QNSAAKVSPA TRSDTVATMM REKGMYFRRE  360
LDRYSLDSED LYSRNAGPQA NFRNKRGQMP ENPYSEVGKI ASKAVYVPAK PARRKGMLVK  420
QSNVEDSPEK TCSIPIPTII VKEPSTSSSG KSSQGSSMEI DPQAPEPPSQ LRPDESLTVS  480
SPFAAAIAGA VRDREKRLEA RRNSPAFLST DLGDEDVGLG PPAPRTRPSM FPEEGDFADE  540
DSAEQLSSPM PSATPREPEN HFVGGAEASA PGEAGRPLNS TSKAQGPESS PAVPSASSGT  600
AGPGNYVHPL TGRLLDPSSP LALALSARDR AMKESQQGPK GEAPKADLNK PLYIDTKMRP  660
SLDAGFPTVT RQNTRGPLRR QETENKYETD LGRDRKGDDK KNMLIDIMDT SQQKSAGLLM  720
VHTVDATKLD NALQEEDEKA EVEMKPDSSP SEVPEGVSET EGALQISAAP EPTTVPGRTI  780
VAVGSMEEAV ILPFRIPPPP LASVDLDEDF IFTEPLPPPL EFANSFDIPD DRAASVPALS  840
DLVKQKKSDT PQSPSLNSSQ PTNSADSKKP ASLSNCLPAS FLPPPESFDA VADSGIEEVD  900
SRSSSDHHLE TTSTISTVSS ISTLSSEGGE NVDTCTVYAD GQAFMVDKPP VPPKPKMKPI  960
IHKSNALYQD ALVEEDVDSF VIPPPAPPPP PGSAQPGMAK VLQPRTSKLW GDVTEIKSPI  1020
LSGPKANVIS ELNSILQQMN REKLAKPGEG LDSPMGAKSA SLAPRSPEIM STISGTRSTT  1080
VTFTVRPGTS QPITLQSRPP DYESRTSGTR RAPSPVVSPT EMNKETLPAP LSAATASPSP  1140
ALSDVFSLPS QPPSGDLFGL NPAGRSRSPS PSILQQPISN KPFTTKPVHL WTKPDVADWL  1200
ESLNLGEHKE AFMDNEIDGS HLPNLQKEDL IDLGVTRVGH RMNIERALKQ LLDR
```

SEQ ID NO:166 PEZ4 DNA SEQUENCE

Nucleic Acid Accession #:   NM_000024
Coding sequence: 220-1461 (underlined sequences correspond to start and stop codons)

```
       1     11    21    31    41    51
       |     |     |     |     |     |
ACTGCGAAGC GGCTTCTTCA GAGCACGGGC TGGAACTGGC AGGCACCGCG AGCCCCTAGC  60
ACCCGACAAG CTGAGTGTGC AGGACGAGTC CCCACCACAC CCACACCACA GCCGCTGAAT  120
GAGGCTTCCA GGCGTCCGCT CGCGGCCCGC AGAGCCCCGC CGTGGGTCCG CCCGCTGAGG  180
CGCCCCCAGC CAGTGCGCTT ACCTGCCAGA CTGCGCGCCCA TGGGGCAACC CGGGAACGGC  240
AGCGCCTTCT TGCTGGCACC CAATAGAAGC CATGCGCCGG ACCACGACGT CACGCAGCAA  300
AGGGACGAGG TGTGGGTGGT GGGCATGGGC ATCGTCATGT CTCTCATCGT CCTGGCCATC  360
GTGTTTGGCA ATGTGCTGGT CATCACAGCC ATTGCCAAGT TCGAGCGTCT GCAGACGGTC  420
ACCAACTACT TCATCACTTC ACTGGCCTGT GCTGATCTGG TCATGGGCCT GGCAGTGGTG  480
CCCTTTGGGG CCGCCCATAT TCTTATGAAA ATGTGGACTT TTGGCAACTT CTGGTGCGAG  540
TTTTGGACTT CCATTGATGT GCTGTGCGTC ACGGCCAGCA TTGAGACCCT GTGCGTGATC  600
GCAGTGGATC GCTACTTTGC CATTACTTCA CCTTTCAAGT ACCAGAGCCT GCTGACCAAG  660
AATAAGGCCC GGGTGATCAT TCTGATGGTG TGGATTGTGT CAGGCCTTAC CTCCTTCTTG  720
CCCATTCAGA TGCACTGGTA CCGGGCCACC CACCAGGAAG CCATCAACTG CTATGCCAAT  780
GAGACCTGCT GTGACTTCTT CACGAACCAA GCCTATGCCA TTGCCTCTTC CATCGTGTCC  840
TTCTACGTTC CCCTGGTGAT CATGGTCTTC GTCTACTCCA GGGTCTTTCA GGAGGCCAAA  900
AGGCAGCTCC AGAAGATTGA CAAATCTGAG GGCCGCTTCC ATGTCCAGAA CCTTAGCCAG  960
GTGGAGCAGG ATGGGCGGAC GGGGCATGGA CTCCGCAGAT CTTCCAAGTT CTGCTTGAAG  1020
GAGCACAAAG CCCTCAAGAC GTTAGGCATC ATCATGGGCA CTTTCACCCT CTGCTGGCTG  1080
CCCTTCTTCA TCGTTAACAT TGTGCATGTG ATCCAGGATA ACCTCATCCG TAAGGAAGTT  1140
TACATCCTCC TAAATTGGAT AGGCTATGTC AATTCTGGTT TCAATCCCCT TATCTACTGC  1200
CGGAGCCCAG ATTTCAGGAT TGCCTTCCAG GAGCTTCTGT GCCTGCGCAG GTCTTCTTTG  1260
AAGGCCTATG GGAATGGCTA CTCCAGCAAC GGCAACACAG GGGAGCAGAG TGGATATCAC  1320
GTGGAACAGG AGAAAGAAAA TAAACTGCTG TGTGAAGACC TCCCAGGCAC GGAAGACTTT  1380
GTGGGCCATC AAGGTACTGT GCCTAGCGAT AACATTGATT CACAAGGGAG GAATTGTAGT  1440
ACAAATGACT CACTGCTGTA AAGCAGTTTT TCTACTTTTA AAGACCCCCC CCCCCCCAAC  1500
AGAACACTAA ACAGACTATT TAACTTGAGG GTAATAAACT TAGAATAAAA TTGTAAAAAT  1560
TGTATAGAGA TATGCAGAAG GAAGGGCATC CTTCTGCCTT TTTTATTTTT TTAAGCTGTA  1620
AAAAGAGAGA AAACTTATTT GAGTGATTAT TTGTTATTTG TACAGTTCAG TTCCTCTTTG  1680
CATGGAATTT GTAAGTTTAT GTCTAAAGAG CTTTAGTCCT AGAGGACCTG AGTCTGCTAT  1740
ATTTTCATGA CTTTTCCATG TATCTACCTC ACTATTCAAG TATTAGGGGT AATATATTGC  1800
TGCTGGTAAT TTGTATCTGA AGGAGATTTT CCTTCCTACA CCCTTGGACT TGAGGATTTT  1860
GAGTATCTCG GACCTTTCAG CTGTGAACAT GGACTCTTCC CCCACTCCTC TTATTTGCTC  1920
ACACGGGGTA TTTTAGGCAG GGATTTGAGG AGCAGCTTCA GTTGTTTTCC CGAGCAAAGG  1980
TCTAAAGTTT ACAGTAAAATA AAATGTTTGA CCATG
```

SEQ ID NO:167 PEZ4 Protein sequence:
Protein Accession #:   NP_000015.1

```
       1     11    21    31    41    51
       |     |     |     |     |     |
```

368

```
 |    |    |    |    |    |
MGQPGNGSAF LLAPNRSHAP DHDVTQQRDE VWVVGMGIVM SLIVLAIVFG NVLVITAIAK  60
FERLQTVTNY FITSLACADL VMGLAVVPFG AAHILMKMWT FGNFWCEFWT SIDVLCVTAS  120
IETLCVIAVD RYFAITSPFK YQSLLTKNKA RVIILMVWIV SGLTSFLPIQ MHWYRATHQE  180
AINCYANETC CDFFTNQAYA IASSIVSFYV PLVIMVFVYS RVFQEAKRQL QKIDKSEGRF  240
HVQNLSQVEQ DGRTGHGLRR SSKFCLKEHK ALKTLGIIMG TFTLCWLPFF IVNIVHVIQD  300
NLIRKEVYIL LNWIGYVNSG FNPLIYCRSP DFRIAFQELL CLRRSSLKAY GNGYSSNGNT  360
GEQSGYHVEQ EKENKLLCED LPGTEDFVGH QGTVPSDNID SQGRNCSTND SLL
```

<center>SEQ ID NO:168 PEZ1 DNA SEQUENCE</center>

Nucleic Acid Accession #:   NM_004457
Coding sequence: 143-2305   (underlined sequences correspond to start and stop codons)

```
 1     11    21    31    41    51
 |     |     |     |     |     |
GAATTCGTTG TTGGGAAGGA CTGGGGAAAC AGCTGTAACA TTTGCCACCC TCAGAAGCTG   60
CTGGTCCTGT GTCACACCAC CTTAGCCTCT TGATCGAGGA AGATTCTCGC TGAAGTCTGT  120
TAATTCTACT TTTTGAGTAC TTATGAATAA CCACGTGTCT TCAAAACCAT CTACCATGAA  180
GCTAAAACAT ACCATCAACC CTATTCTTTT ATATTTTATA CATTTTCTAA TATCACTTTA  240
TACTATTTTA ACATACATTC CGTTTTATTT TTTCTCCGAG TCAAGACAAG AAAAATCAAA  300
CCGAATTAAA GCAAAGCCTG TAAATTCAAA ACCTGATTCT GCATGACGAT CTGTTAATAG  360
TTTGGATGGT TTGGCTTCAG TATTATACCC TGGATGTGAT ACTTTAGATA AAGTTTTTAC  420
ATATGCAAAA AACAAATTTA AGAACAAAAG ACTCTTGGGA ACACGTGAAG TTTTAAATGA  480
GGAAGATGAA GTACAACCAA ATGGAAAAAT TTTTAAAAAG GTTATTCTTG GACAGTATAA  540
TTGGCTTTCC TATGAAGATG TCTTTGTTCG AGCCTTTAAT TTTGGAAATG GATTACAGAT  600
GTTGGGTCAG AAACCAAAGA CCAACATCGC CATCTTCTGT GAGACCAGGG CCGAGTGGAT  660
GATAGCTGCA CAGGCGTGTT TTATGTATAA TTTTCAGCTT GTTACATTAT ATGCCACTCT  720
AGGAGGTCCA GCCATTGTTC ATGCATTAAA TGAAACAGAG GTGACCAACA TCATTACTAG  780
TAAAGAACTC TTACAAACAA AGTTGAAGGA TATAGTTTCT TTGGTCCCAC GCCTGCGGCA  840
CATCATCACT GTTGATGGAA AGCCACCGAC CTGGTCCGAC TTCCCCAAGG GCATCATTGT  900
GCATACCATG GCTGCAGTGG AGGCCCTGGG AGCCAAGGCC AGCATGGAAA ACCAACCTCA  960
TAGCAAACCA TTGCCCTCAG ATATTGCAGT AATCATGTAC ACAAGTGGAT CCACAGGACT 1020
TCCAAAGGGA GTCATGATCT CACATAGTAA CATTATTGCT GGTATAACTG GGATGGCAGA 1080
AAGGATTCCA GAACTAGGAG AGGAAGATGT CTACATTGGA TATTTGCCTC TGGCCCATGT 1140
TCTAGAATTA AGTGCTGAGC TTGTCTGTCT TTCTCACGGA TGCCGCATTG GTTACTCTTC 1200
ACCACAGACT TTAGCAGATC AGTCTTCAAA AATTAAAAAA GGAAGCAAAG GGGATACATC 1260
CATGTTGAAA CCAACACTGA TGGCAGCAGT TCCGGAAATC ATGGATCGGA TCTACAAAAA 1320
TGTCATGAAT AAAGTCAGTG AAATGAGTAG TTTTCAACGT AATCTGTTTA TTCTGGCCTA 1380
TAATTACAAA ATGGAACAGA TTTCAAAAGG ACGTAATACT CCACTGTGCG CACAGCTTTGT 1440
TTTCCGGAAA GTTCGAAGCT TGCTAGGGGG AAATATTCGT CTCCTGTTGT GTGGTGGCGC 1500
TCCACTTTCT GCAACCACGC AGCGATTCAT GAACATCTGT TTCTGCTGTC CTGTTGGTCA 1560
GGGATACGGG CTCACTGAAT CTGCTGGGGC TGGAACAATT TCCGAAGTGT GGGACTACAA 1620
TACTGGCAGA GTGGGAGCAC CATTAGTTTG CTGTGAAATC AAATTAAAAA ACTGGGAGGA 1680
AGGTGGATAC TTTAATACTG ATAAGCCACA CCCCAGGGGT GAAATTCTTA TTGGGGGCCA 1740
AAGTGTGACA ATGGGGTACT ACAAAAATGA AGCAAAAACA AAAGCTGATT TCTCTGAAGA 1800
TGAAAATGGA CAAAGGTGGC TCTGTACTGG GGATATTGGA GAGTTTGAAC CCGATGGATG 1860
CTTAAAGATT ATTGATCGTA AAAAGGACCT TGTAAAACTA CAGGCAGGGG AATATGTTTC 1920
TCTTGGGAAA GTAGAGGCAG CTTTGAAGAA TCTTCCACTA GTAGATAACA TTTGTGCATA 1980
TGCAAACAGT TATCATTCTT ATGTCATTGG ATTTGTTGTG CCAAATCAAA AGGAACTAAC 2040
TGAACTAGCT CGAAAGAAAG GACTTAAAGG GACTTGGGAG GAGCTGTGTA ACAGTTGTGA 2100
AATGGAAAAT GAGGTACTTA AAGTGCTTTC CGAAGCTGCT ATTTCAGCAA GTCTGGAAAA 2160
GTTTGAAATT CCAGTAAAAA TTCGTTTGAG TCCTGAACCG TGGACCCCTG AAACTGGTTT 2220
GGTGACAGAT GCCTTCAAGC TGAAACGCAA AGAGCTTAAA ACACATTACC AGGCGGACAT 2280
TGAGCGAATG TATGGAAGAA AATAATTATT CTCTTCTGGC ATCAGTTTGC TACAGTGAGC 2340
TCACATCAAA TAGGAAAATA CTTGAAATGC ATGTCTCAAG CTGCAAGGCA AACTCCATTC 2400
CTCATATTAA ACTATTACTT CTCATGACGT CACCATTTTT AACTGACAGG ATTAGTAAAA 2460
CATTAAGACA GCAAACTTGT GTCTGTCTCT TCTTTCATTT TCCCCGCCAC CAACTTACTT 2520
TACCACCTAT GACTGTACTT GTCAGTATGA GAATTTTTCT GAATCATATT GGGGAAGCAG 2580
TGATTTTAAA ACCTCAAGTT TTTAAACATG ATTTATATGT TCTGTATAAT GTTCAGTTTG 2640
TAACTTTTTA AAAGTTTGGA TGTATAGAGG GATAAATAGG AAATATAAGA ATTGGTTATT 2700
TGGGGGCTTT TTTACTTACT GTATTTAAAA ATACAAGGGT ATTGATATGA AATTATGTAA 2760
ATTTCAAATG CTTATGAATC AAATCATTGT TGAACAAAAG ATTTGTTGCT GTGTAATTAT 2820
TGTCTTGTAT GCATTTGAGA GAAATAAATA TACCCATACT TATGTTTTAA GAAGTTGAGA 2880
TCTTGTGAAT ATATGCCTGT CAGTGTCTTC TTTATATATT TATTTTTTAT TAGAAAAAAT 2940
GAAGTTTGGT TGGTGATGCA TGAAACAAAA TAGCAAGAGA GGGTTATAGT TTAAATAGTAA 3000
GGGAGATAAC ACAGCATGTG TAGCACCAGT TGATAATTGG TCTCTAGTAG CTTACTGTCA 3060
AAATGTTCAA TGAAGTCTTC TGTTCATCTG TTGAAACTAG GAAAATACCC AAACTTAAAT 3120
GGAAGAATTC TGAAAGAGAG GATAGAATTT AAAGAACAAG AGTATATAAA GTTATTCTTT 3180
GAATATTTCG TTGACTATAT GTACATTGAG TTATCTATAT TTGTAAACAA ATTAGTCATG 3240
GAAAATTATT CTATTCCAAA GTCTCCTTTT AGTCTAGATA ATCATTATTT CATTTTAAAA 3300
TTAGTGTTTT TCATAGTTTG CACTGATGCG TGTATGGATG TGTGTGAGTC AGTGGTAGCT 3360
TATTTAAAAA GCACCTTATC CTTTCTCCCA TAACCTTTGT ACACTAAAAA ATGAAAGAAT 3420
TTAGAATGTA TTTGATGATA GCATTCTCAC TAAGACACAT GAGAATTTAA CTTTATAACC 3480
GCGTGAGTTA AGATTTAATT CATAGGTTTT GATGTGCATTG TTGAAGTTAT TTGTAATTCA 3540
GAAACCTTGC TTGTGTGATA CATAGTAAGT CTCTTCATTT ATTACTGCTT GCCTGTTGTT 3600
```

<center>369</center>

ATATCTGGAT TATCAAAAGC AATAGTGCAC CAATTAAGAT GTGCTCAAAT CAGGACTTAA 3660
ATCATAGGCA CCACATTTTT CATGTCAGAC TAGTTACTTT GTTGATTCTC AGTTACTGTA 3720
GGCATCAAAA GGCAAAAATC A

SEQ ID NO:169 PEZ1 Protein sequence:
Protein Accession #:   NP_004448.1

```
1   11   21   31   41   51
|   |    |    |    |    |
MNNHVSSKPS TMKLKHTINP ILLYFIHFLI SLYTILTYIP FYFFSESRQE KSNRIKAKPV   60
NSKPDSAYRS VNSLDGLASV LYPGCDTLDK VFTYAKNKFK NKRLLGTREV LNEEDEVQPN  120
GKIFKKVILG QYNWLSYEDV FVRAFNFGNG LQMLGQKPKT NIAIFCETRA EWMIAAQACF  180
MYNFQLVTLY ATLGGPAIVH ALNETEVTNI ITSKELLQTK LKDIVSLVPR LRHIITVDGK  240
PPTWSDFPKG IIVHTMAAVE ALGAKASMEN QPHSKPLPSD IAVIMYTSGS TGLPKGVMIS  300
HSNIIAGITG MAERIPELGE EDVYIGYLPL AHVLELSAEL VCLSHGCRIG YSSPQTLADQ  360
SSKIKKGSKG DTSMLKPTLM AAVPEIMDRI YKNVMNKVSE MSSFQRNLFI LAYNYKMEQI  420
SKGRNTPLCD SFVFRKVRSL LGGNIRLLLC GGAPLSATTQ RFMNICFCCP VGQGYGLTES  480
AGAGTISEVW DYNTGRVGAP LVCCEIKLKN WEEGGYFNTD KPHPRGEILI GGQSVTMGYY  540
KNEAKTKADF SEDENGQRWL CTGDIGEFEP DGCLKIIDRK KDLVKLQAGE YVSLGKVEAA  600
LKNLPLVDNI CAYANSYHSY VIGFVVPNQK ELTELARKKG LKGTWEELCN SCEMENEVLK  660
VLSEAAISAS LEKFEIPVKI RLSPEPWTPE TGLVTDAFKL KRKELKTHYQ ADIERMYGRK
```

SEQ ID NO:170 PCQ7 DNA SEQUENCE
Nucleic Acid Accession #:   none found
        Coding sequence: 38-1075(underlined sequence corresponds to start and stop codon)

```
1          11         21         31         41         51
|          |          |          |          |          |
AGCAACGACG CCGGGCAGCG GGAGCGGCGG CCGCGCCATG TGGCTGCTGG GGCCGCTGTG    60
CCTGCTGCTG AGCAGCGCCG CGGAGAGCCA GCTGCTCCCC GGGAACAACT TCACCAATGA   120
GTGCAACATA CCAGGCAACT TCATGTGCAG CAATGGACGG TGCATCCCGG GCGCCTGGCA   180
GTGTGACGGG CTGCCTGACT GCTTCGACAA GAGTGATGAG AAGGAGTGCC CCAAGGCTAA   240
GTCGAAATGT GGCCCAACCT TCTTCCCCTG TGCCAGCGGC ATCCATTGCA TCATTGGTCG   300
CTTCCGGTGC AATGGGTTTG AGGACTGTCC CGATGGCAGC GATGAAGAGA ACTGCACAGC   360
AAACCCTCTG CTTTGCTCCA CCGCCCGCTA CCACTGCAAG AACGGCCTCT GTATTGACAA   420
GAGCTTCATC TGCGATGGAC AGAATAACTG TCAAGACAAC AGTGATGAGG AAAGCTGTGA   480
AAGTTCTCAA GAACCCGGCA GTGGGCAGGT GTTTGTGACT TCAGAGAACC AACTTGTGTA   540
TTACCCCAGC ATCACCTATG CCATCATCGG CAGCTCCGTC ATTTTTGTGC TGGTGGTGGC   600
CCTGCTGGCA CTGGTCTTGC ACCACCAGCG GAAGCGGAAC AACCTCATGA CGCTGCCCGT   660
GCACCGGCTG CAGCACCCTG TGCTGCTGTC CCGCCTGGTG GTCCTGGACC ACCCCCACCA   720
CTGCAACGTC ACCTACAACG TCAATAATGG CATCCAGTAT GTGGCCAGCC AGGCGGAGCA   780
GAATGCGTCG GAAGTAGGCT CCCCCACCCTC CTACTCCGAG GCCTTGCTGG ACCAGAGGCC   840
TGCGTGGTAT GACCTTCCTC CACCGCCCTA CTCTTCTGAC ACGGAATCTC TGAACCAAGC   900
CGACCTGCCC CCCTACCGCT CCCGGTCCGG GAGTGCCAAC AGTGCCAGCT CCCAGGCAGC   960
CAGCAGCCTC CTGAGCGTGG AAGACACCAG CCACAGCCCG GGGCAGCCTG GCCCCCAGGA  1020
GGGCACTGCT GAGCCCAGGG ACTCTGAGCC CAGCCAGGGC ACTGAAGAAG TATAAGTCCC  1080
AGTTATTCCA AAGTCCATAT GGGTTAATCT GCTCTGACTT GTTGCCATTC TAACAATTTG  1140
TGCTCATGGG AAGCTCTTTA AGCACCTGTA AGGATGTCTC AAGTTACAGT TTGGGATATT  1200
AACTATCTCT GCATTCCCCT CCTCCCCCAG ACTTCAGAGA TGTTTTTCTG GCGTCTCAGT  1260
TGACATGATC TGTTGTGCGT CTTTTTCTGTC AGGTCACTCT TCCCTTGGGA CCCGAGATCA  1320
CACCCTCATT TTTCACATTA TTCTGTTTCT GTTGGAGAGA CAGCATATAA AACAGTATTG  1380
AAATAGGCTG GGAGAGAGCA ATGTTTCTGT GCTATATTGG ATGCTCAGAA GTGCAGGAGA  1440
CGCTGGACCC AATTCTCTCT GCTGGGTAGT TACCTTATAG CATTTGGGGA TTTGGGTTAG  1500
ATGATCTAAC CAGGAGGCCA TCACTGGATG GTCACCCCCC CAAAAAAATT CCATTTGAGC  1560
ATCAAAACCT GCTTTGCACA ATCCTATTTG ATGCCCCCAG TTCAGCAGAG TCAGTGGCCA  1620
AAGAAAACTT TGGACGTGAG TAACACCCTT CAGCAGTCGC AACGTTATTT TGGTTTTGTG  1680
AAGGACTCTG AAACCATCTA CCCTGTATAA ATTCTGGCTT TAGAAATTTG CCCAAGAATG  1740
CTCATTCTGA GAGCTTTCCT CAGCAGCATA TATCATCAGC CTCATCCTAA AATAGGCAGG  1800
GAGCCCCTCC CATGAGTTTA TCCAAGTTCT CAGCTCCTAA AATGCAGGCT GCCAAGACCC  1860
TACACCTGCC CTGGCTCTAC AGCCACTTAC CTGGTTTCTG GACTGTCACC CTCCCAGCTG  1920
ACCTGCCCGT AGCCAAGGAA TGAGGACCTA ACTTGAGTTG GCCCAAAGTC TGACCTGGCT  1980
GTATGTCCCT GTGGCCCACA CCCAGCCTGT CTTGCTCATT CATGCAGCCT CAACACTGGC  2040
CTCCAAAGTT CCCTTAACAC TTGCAAAGTC CTTTTTACCT GTGCATTTGG ACTTGAGGAC  2100
ACTGGTTTCT ATCACAGGTG AGAGCCATGT TCAATACCTC CAGCAAGCTC TCCTGGCTCC  2160
CTGCACTGTG CACGCTCCTC TTCCCAAGGT CCCAATACCA GCACCTCTAG TTAGAGTTAG  2220
GGTCAGGGTC AGGCCTCTCC CAACATCCCA GTAGTTTCTC CTCTGAGACA CATGGGCAAG  2280
AGACAAATTTG GAGTCAAGAT TTTCCATTTG GATCTATTTT AAATCTTTTA GAAATGCATT  2340
TGAAACAGTG TGTTTGTTTT TTCCCTTCTA GTTAAGGGAC TATTTATATG TGTATAGGAA  2400
AGCTGTCTCT TTTTTTGTTT TTCCTTTAAC AAGGTCCAAA GAAAGATGCA AAAGGAGATC  2460
ACACCCTTGC CCCGCTGAGC CCCGTGATAA CAAGTCACTC CAGACTAACC TGTGTGCCAG  2520
ACATTTGTGC ATTGTTGCAC TTTGAGGTTA TTATTTATCA AGTTCTTGAA GGAAGCAGAA  2580
AGAGGGACTC CTCTCTCCCT CCGTGTATAG TCTCTATGTT TGTGCTAGTT TTTCTTTTTT  2640
TTCTCTGTGT CCAGTCAGCC ACAGGGCCCG CCTCCCTGCA GGAATAAGGG GTAAAACGTT  2700
AGGTGTTGTT TGGCAAGAAA CCACACTGAC TGATGAGGGG TAAAATGGAA CCAGGTAGAG  2760
CCACTCCGGG CAGCTGTCAC CCATTCAGAA CTTCTTTCCG CAGCTGAAGA AATGTTCAGT  2820
AACCTGTTTG ACGCTAATTA AAACAGAGCC TGCAGGAAGT GGGGCTAAAG TGGCATTCAG  2880
TGATCCTGTT CTGTAGACTT TTCTTTCTTT TTTTAACCAA ATCCAAAGGA TGTTACAGAA  2940
```

370

```
AAGCTAGCCA CTGGTATTTT GTTTTGTTTA AAAAAAAAAA GAAAGAAAGA AAGAAAGAAA   3000
AACGGAAAGG AACCTAGCTG CCTGTATCTT TCATTTTTAA AATAGCACTT GAGTTATTTT   3060
CTGAGTAATC CAATAAAGAA CTTTTGATGA CAGCCAGAAT GTGTTAGAAC TCTGGCTGAA   3120
CATTTCATCT CCTGTGAGTC AGAAGGGCTT TATTTCTCCC TTTGATGGGG CCCCTTCTTC   3180
TTTCTGGTGC TCTGGAAGTT GTTTAGAGGA AAGAATTCTA ATTTTAATTA ATTGCGCAGT   3240
GAGTTAATCT CACTCGCTTT TCTGCTTCCA GGCATCTTAG GAAAAACAAA TGGTTTTAGT   3300
AGATAAGGGA TGCCTACTAA TGCTTTTTTA AAACAAACAG GGACATTTTT ATTATAGATT   3360
TGATTTTTTT AATGAATGTT TTTAAAAATA TATAAATAGG ACACCAAAGC GGCAGGGTTT   3420
TTTTTGGGGG GAGGGGGTTT GTTTTCCAAC TCAAGATGGC ACATTAGTGG CCAGCAATAT   3480
TTTTTAACTC ATTCCAACCA GGAAGCTTTT TTATACATTG CCTAAATCTA CGCCAACCAG   3540
AAAATAGTCT CATCTCTTTT TTTCTCAAAT GAGATCCGTG TTTTATTTTA GCATTAAATT   3600
AGTTACACTG TGATGACTGG CCTATTACCT GACTCAGCTC CCTCTACCTT GAAATTGACA   3660
TTTTTAAAAA ATGCAACTAA GTGGTTAATA GTGTGTGACG CTCAAAGTTA ATGTAAACTG   3720
GAAAGGTTGT GTGTCGTTGC TTTTTGTGTT TTGGTTAGGC TTGGTTTTGT TTTTTAATTT   3780
TTATACTTTC TAATAAATTT GCAGTTTCAT TCTTTCTGTT TGTGCAAAWG GWMCTAMARM   3840
AAMMAAAAAC AWYWTTGGGG GGGCTTGGGC CTCCGGAAAAA GTTTTTAACA CCACTTCGGG   3900
TGGGGCGGCG GGGCCCACGT AGGTACGGCG ACCACGCGGG CCCAAACGGG ACCCCAGAAG   3960
GAAACCCTGG CCAAGAAAAA GGTGGCGAGA ATTCTCCACA CCAGAAAAAA ACGCGCCGGG   4020
GGAAACCGCA GAGTGTTGCG TAAACCACAC CCGAAGAGAG AACTCAGAAG CACACAAGCG   4080
GGACTCAACC AGGAGGACCC AAGGGAACCC GATAGAGTAC G
```

SEQ ID NO:171 PCQ7 Protein sequence:
Protein Accession #:   none found

```
1          11         21         31         41         51
|          |          |          |          |          |
MWLLGPLCLL LSSAAESQLL PGNNFTNECN IPGNFMCSNG RCIPGAWQCD GLPDCFDKSD   60
EKECPKAKSK CGPTFFPCAS GIHCIIGRFR CNGFEDCPDG SDEENCTANP LLCSTARYHC   120
KNGLCIDKSF ICDGQNNCQD NSDEESCESS QEPGSGQVFV TSENQLVYYP SITYAIIGSS   180
VIFVLVVALL ALVLHHQRKR NNLMTLPVHR LQHPVLLSRL VVLDHPHHCN VTYNVNNGIQ   240
YVASQAEQNA SEVGSPPSYS EALLDQRPAW YDLPPPPYSS DTESLNQADL PPYRSRSGSA   300
NSASSQAASS LLSVEDTSHS PGQPGPQEGT AEPRDSEPSQ GTEEV
```

                                      SEQ ID NO:172 PEL3 DNA SEQUENCE
Nucleic Acid Accession #:   NM_005656.1
    Coding sequence: 57-1535 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
GTCATATTGA ACATTCCAGA TACCTATCAT TACTCGATGC TGTTGATAAC AGCAAGATGG   60
CTTTGAACTC AGGGTCACCA CCAGCTATTG GACCTTACTA TGAAAACCAT GGATACCAAC   120
CGGAAAACCC CTATCCCGCA CAGCCCACTG TGGTCCCCAC TGTCTACGAG GTGCATCCGG   180
CTCAGTACTA CCCGTCCCCC GTGCCCCAGT ACGCCCCGAG GGTCCTGACG CAGGCTTCCA   240
ACCCCGTCGT CTGCACGCAG CCCAAATCCC CATCCGGGAC AGTGTGCACC TCAAAGACTA   300
AGAAAGCACT GTGTCATCACC TTGACCCTGG GGACCTTCCT CGTGGGAGCT GCGCTGGCCG   360
CTGGCCTACT CTGGAAGTTC ATGGGCAGCA AGTGCTCCAA CTCTCGGGATA GAGTGCGACT   420
CCTCAGGTAC CTGCATCAAC CCCTCTAACT GGTGTGATGG CGTGTCACAC TGCCCCGGCG   480
GGGGACGA GAATCGGTGT GTTCGCTCT ACGGACCAAA CTTCATCCTT CAGATGTACT   540
CATCTCAGAG GAAGTCCTGG CACCCTGTGT GCCAAGACGA CTGGAACGAG AACTACGGGC   600
GGGCGGCCTG CAGGGACATG GGCTATAAGA ATAATTTTTA CTCTAGCCAA GGAATAGTGG   660
ATGACAGCGG ATCCACCAGC TTTATGAAAC TGAACACAAG TGCCGGCAAT GTCGATATCT   720
ATAAAAAACT GTACCACAGT GATGCCTGTT CTTCAAAAGC AGTGGTTTCT TTACGCTGTT   780
TAGCCTGCGG GGTCAACTTG AACTCAAGCC GCCAGAGCAG GATCGTGGGC GGTGAGAGCG   840
CGCTCCCGGG GGCCTGGCCC TGGCAGGTCA GCCTGCACGT CCAGAACGTC CACGTGTGCG   900
GAGGCTCCAT CATCCACCCC GAGTGGATCG TGACAGCCGC CCACTGCGTG GAAAAACCTC   960
TTAACAATCC ATGGCATTGG ACGGCATTTG CGGGGATTTT GAGACAATCT TTCATGTTCT   1020
ATGGAGCCGG ATACCAAGTA CAAAAAGTGA TTTCTCATCC AAATTATGAC TCCAAGACCA   1080
AGAACAATGA CATTGCGCTG ATGAAGCTGC AGAAGCCTCT GACTTTCAAC GACCTAGTGA   1140
AACCAGTGTG TCTGCCCAAC CCAGGCATGA TGCTGCAGCC AGAACAGCTC TGCTGGATTT   1200
CCGGGTGGGG GGCCACCGAG GAGAAGGGA AGACCTCAGA AGTGCTGAAC GCTGCCAAGG   1260
TGCTTCTCAT TGAGACACAG AGATGCAACA GCAGATATGT CTATGACAAC CTGATCACAC   1320
CAGCCATGAT CTGTGCCGGC TTCCTGCAGG GGAACGTCGA TTCTTGCCAG GGTGACAGTG   1380
GAGGGCCTCT GGTCACTTCG AACAACAATA TCTGGTGGCT GATAGGGGAT ACAAGCTGGG   1440
GTTCTGGCTG TGCCAAAGCT TACAGACCAG GAGTGTACGG GAATGTGATG GTATTCACGG   1500
ACTGGATTTA TCGACAAATG AAGGCAAACG GCTAATCCAC ATGGTCTTCG TCCTTGACGT   1560
CGTTTTACAA GAAAACAATG GGGCTGGTTT TGCTTCCCCG TGCATGATTT ACTCTTAGAG   1620
ATGATTCAGA GGTCACTTCA TTTTTATTAA ACAGTGAACT TGTCTGGCTT TGGCACTCTC   1680
TGCCATACTG TGCAGGCTGC AGTGGCTCCC CTGCCCAGCC TGCTCTCCCT AACCCCTTGT   1740
CCGCAAGGGG TGATGGCCGG CTGGTTGTGG GCACTGGCGG TCAATTGTGG AAGGAAGAGG   1800
GTTGGAGGCT GCCCCCATTG AGATCTTCCT GCTGAGTCCT TTCCAGGGGC CAATTTTGGA   1860
TGAGCATGGA GCTGTCACTT CTCAGCTGCT GGATGACTTG AGATGAAAAA GGAGAGACAT   1920
GGAAAGGGAG ACAGCCAGGT GGCACCTGCA GCGGCTGCCC TCTGGGGCCA CTTGGTAGTG   1980
TCCCCAGCCT ACTTCACAAG GGGATTTTGC TGATGGGTTC TTAGAGCCTT AGCAGCCCTG   2040
GATGGTGGCC AGAAATAAAG GGACCAGCCC TTCATGGGTG GTGACGTGGT AGTCACTTGT   2100
AAGGGGAACA GAAACATTTT TGTTCTTATG GGGTGAGAAT ATAGACAGTG CCCTTGGTGC   2160
```

```
     GAGGGAAGCA ATTGAAAAGG AACTTGCCCT GAGCACTCCT GGTGCAGGTC TCCACCTGCA 2220
     CATTGGGTGG GGCTCCTGGG AGGGAGACTC AGCCTTCCTC CTCATCCTCC CTGACCCTGC 2280
     TCCTAGCACC CTGGAGAGTG AATGCCCCTT GGTCCCTGGC AGGGCGCCAA GTTTGGCACC 2340
  5  ATGTCGGCCT CTTCAGGCCT GATAGTCATT GGAAATTGAG GTCCATGGGG GAAATCAAGG 2400
     ATGCTCAGTT TAAGGTACAC TGTTTCCATG TTATGTTTCT ACACATTGAT GGTGGTGACC 2460
     CTGAGTTCAA AGCCATCTT
```

<u>SEQ ID NO:173 PEL3 Protein sequence:</u>
Protein Accession #:   NP_005647.1

```
     1          11         21         31         41         51
     |          |          |          |          |          |
     MALNSGSPPA IGPYYENHGY QPENPYPAQP TVVPTVYEVH PAQYYPSPVP QYAPRVLTQA   60
     SNPVVCTQPK SPSGTVCTSK TKKALCITLT LGTFLVGAAL AAGLLWKFMG SKCSNSGIEC  120
     DSSGTCINPS NWCDGVSHCP GGEDENRCVR LYGPNFILQM YSSQRKSWHP VCQDDWNENY  180
     GRAACRDMGY KNNFYSSQGI VDDSGSTSFM KLNTSAGNVD IYKKLYHSDA CSSKAVVSLR  240
     CLACGVNLNS SRQSRIVGGE SALPGAWPWQ VSLHVQNVHV CGGSIITPEW IVTAAHCVEK  300
     PLNNPWHWTA FAGILRQSFM FYGAGYQVQK VISHPNYDSK TKNNDIALMK LQKPLTFNDL  360
     VKPVCLPNPG MMLQPEQLCW ISGWGATEEK GKTSEVLNAA KVLLIETQRC NSRYVYDNLI  420
     TPAMICAGFL QGNVDSCQGD SGGPLVTSNN NIWWLIGDTS WGSGCAKAYR PGVYGNVMVF  480
     TDWIYRQMKA NG
```

<center>SEQ ID NO:174 PBJ4 DNA SEQUENCE</center>
Nucleic Acid Accession #:   AI694767
                          Coding sequence: 130-1086 (underlined sequences correspond to start and stop codons)

```
     1          11         21         31         41         51
     |          |          |          |          |          |
     CAGAGAGGCT GTATTTCAGT GCAGCCTGCC AGACCTCTTC TGGAGGAAGA CTGGACAAAG   60
     GGGGTCACAC ATTCCTTCCA TACGGTTGAG CCTCTACCTG CCTGGTGCTG GTCACAGTTC  120
     AGCTTCTTCA TGATGGTGGA TCCCAATGGC AATGAATCCA GTGCTACATA CTTCATCCTA  180
     ATAGGCCTCC CTGGTTTAGA AGAGGCTCAG TTCTGGTTGG CCTTCCCATT GTGCTCCCTC  240
     TACCTTATTG CTGTGCTAGG TAACTTGACA ATCATCTACA TTGTGCGGAC TGAGCACAGC  300
     CTGCATGAGC CCATGTATAT ATTTCTTTGC ATGCTTTCAG GCATTGACAT CCTCATCTCC  360
     ACCTCATCCA TGCCCAAAAT GCTGGCCATC TTCTGGTTCA ATTCACTAC CATCCAGTTT  420
     GATGCTTGTC TGCTACAGAT GTTTGCCATC CACTCCTTAT CTGGCATGGA ATCCACAGTG  480
     CTGCTGGCCA TGGCTTTTGA CCGCTATGTG GCCATCTGTC ACCCACTGCG CCATGCCACA  540
     GTACTTACGT TGCCTCGTGT CACCAAAATT GGTGTGGCTG CTGTGGTGCG GGGGGCTGCA  600
     CTGATGGCAC CCCTTCCTGT CTTCATCAAG CAGCTGCCCT TCTGCCGCTC CAATATCCTT  660
     TCCCATTCCT ACTGCCTACA CCAAGATGTC ATGAAGCTGG CCTGTGATGA TATCCGGGTC  720
     AATGTCGTCT ATGGCCTTAT CGTCATCATC TCCGCCATTG GCCTGGACTC ACTTCTCATC  780
     TCCTTCTCAT ATCTGCTTAT TCTTAAGACT GTGTTGGGCT TGACACGTGA AGCCCAGGCC  840
     AAGGCATTTG GCACTTGCGT CTCTCATGTG TGTGCTGTGT TCATATTCTA TGTACCTTTC  900
     ATTGGATTGT CCATGGTGCA TCGCTTTAGC AAGCGGCGTG ACTCTCCACT GCCCGTCATC  960
     TTGGCCAATA TCTATCTGCT GGTTCCTCCT GTGCTCAACC CAATTGTCTA TGGAGTGAAG 1020
     ACAAAGGAGA TTCGACAGCG CATCCTTCGA CTTTTCCATG TGGCCACACA CGCTTCAGAG 1080
     CCCTAGGTGT CAGTGATCAA ACTTCTTTTC CATTCAGAGT CCTCTGATTC AGATTTTAAT 1140
     GTTAACATTT TGGAAGACAG TATTCAGAAA AAAAATTTCC TTAATAAAAA TACAACTCAG 1200
     ATCCTTCAAA TATGAAACTG GTTGGGGAAT CTCCATTTTT TCAATATTAT TTTCTTCTTT 1260
     GTTTTCTTGC TACATATAAT TATTAATACC CTGACTAGGT TGTGGTTGGA GGGTTATTAC 1320
     TTTTCATTTT ACCATGCAGT CCAAATCTAA ACTGCTTCTA CTGATGGTTT ACAGCATTCT 1380
     GAGATAAGAA TGGTACATCT AGAGAACATT TGCCAAAGGC CTAAGCACAG CAAAGGAAAA 1440
     TAAACACAGA ATATAATAAA ATGAGATAAT CTAGCTTAAA ACTATAACTT CCTCTTCAGA 1500
     ACTCCCAACC ACATTGGATC TCAGAAAAAT ACTGTCTTCA AAATGACTTC TACAGAGAAG 1560
     AAATAATTTT TCCTCTGGAC ACTAGCACTT AAGGGGAAGA TTGGAAGTAA AGCCTTGAAA 1620
     AGAGTACATT TACCTACGTT AATGAAAGTT GACACACTGT TCTGAGAGTT TTCACAGCAT 1680
     ATGGACCCTG TTTTTCCTAT TTAATTTTCT TATCAACCCT TTAATTAGGC AAAGATATTA 1740
     TTAGTACCCT CATTGTAGCC ATGGGAAAAT TGATGTTCAG TGGGGGATCAG TGAATTAAAT 1800
     GGGGTCATAC AAGTATAAAA ATTAAAAAAA AAAGACTTCA TGCCCAATCT CATATGATGT 1860
     GGAAGAACTG TTAAAGAGAC CAACAGGGTA GTGGGTTAGA GATTTCCAGA GTCTTACATT 1920
     TTCTARAGGA GGTATTTAAT TTCTTCTCAC TCATCCAGTG TTGTATTTAG GAATTTCCTG 1980
     GCAACAGAAC TCATGGCTTT AATCCCACTA GCTATTGCTT ATTGTCCTGG TCCAATTGCC 2040
     AATTACCTGT GTCTTGGAAG AAGTGATTTC TAGGTTCACC ATTATGGAAG ATTCTTATTC 2100
     AGAAAGTCTG CATAGGGCTT ATAGCAAGTT ATTTATTTTT AAAAGTTCCA TAGGTGTTTC 2160
     TGATAGGCAG TGAGGTTAGG GAGCCACCAG TTATGATGGG AAGTATGGAA TGGCAGGTGT 2220
     TGAAGATAAC ATTGGCCTTT TGAGTGTGAC TCGTAGCTGG AAAGTGAGGG AATCTTCAGG 2280
     ACCATGCTTT ATTTGGGGCT TTGTGCAGTA TGGAACAGGG ACTTTGAGAC CGGGAAAGCA 2340
     ATCTGACTTA GGCATGGGAA TCAGGCATTT TTGCTTCTGA GGGGCTATTA CCAAGGGTTA 2400
     ATAGGTTTCA TCTTCAACAG GATATGACAA CAGTCTTAAC CAAGAAACTC AAATTACATA 2460
     TACTAAAACA TGTGATCATA TATGTGGTAA GTTTCATTTT CTTTTTCAAT CCTCAGGTTC 2520
     CCTGATATGG ATTCCTATNA CATGCTTTCA TCCCCTTTTG TAATGGATAT CATATTTGGA 2580
     AATGCCTATT TAATACTTGT ATTTGCTGCT GGACTGTAAG CCCATGAGGG CACTGTTTAT 2640
     TATTGAATGT CATCTCTGTT CATCATTGAC TGCTCTTTGC TCATCATTGA ATCCCCCAGC 2700
     AAAGTGCCTA GAACATAATA GTGCTTATGC TTGACACCGG TTATTTTTCA TCAAACCTGA 2760
     TTCCTTCTGT GCTGAACACA TAGCCAGGCA ATTTTCCAGC CTTCTTTGAG TTGGGTATTA 2820
     TTAAATTTTA GCCATTACTT CCAATGTGAG TGGAAGTGAC ATGTGCAATT TTTATACCTG 2880
     GCTCATAAAA CCCTCCCATG TGCAGCCTTT CATGTTGACA TTAAATGTGA CTTGGGAAGC 2940
```

<center>372</center>

```
TATGTGTTAC ACAGAGTTAA TTAACCNGAA AGGCCTGGNA ATTTTTTGNN AANNAAACTG   3000
TGGCCNNGAG GCCCNCAACC CTTTTTNNNA ATTTGGCAAN NTCCCACTTT GTANTTTGGT   3060
AAGGAGGCCA GTTGGATAAG TGAAAAATAA AGTACTATTG TGTC
```

5

SEQ ID NO:175 PBJ4 PROTEIN SEQUENCE
Protein Accession #:   not available, cloned at Eos

10

```
1          11         21         31         41         51
|          |          |          |          |          |
MVDPNGNESS ATYFILIGLP GLEEAQFWLA FPLCSLYLIA VLGNLTIIYI VRTEHSLHEP    60
MYIFLCMLSG IDILISTSSM PKMLAIFWFN STTIQFDACL LQMFAIHSLS GMESTVLLAM   120
AFDRYVAICH PLRHATVLTL PRVTKIGVAA VVRGAALMAP LPVFIKQLPF CRSNILSHSY   180
CLHQDVMKLA CDDIRVNVVY GLIVIISAIG LDSLLISFSY LLILKTVLGL TREAQAKAFG   240
TCVSHVCAVF IFYVPFIGLS MVHRFSKRRD SPLPVILANI YLLVPPVLNP IVYGVKTKEI   300
RQRILRLFHV ATHASEP
```

20

SEQ ID NO:176 PM72 DNA SEQUENCE
Nucleic Acid Accession #:   NM_004624.1
Coding sequence: 57-1544 (underlined sequences correspond to start and stop codons)

```
TCGGAGCCTG CGGAGGGTGG TGGTGGTGGT GGTGGTGGCC CTCGCCCGCC TCACTCATGC    60
CTCCTCCTCC TCTGCTCTCG CTCAGGCGCC TCGGTGGCGG TTGGTCGGCG GTTACGCGGC   120
TGGTGGTCGC GGCGGCCGGG GCTCGCTCTC GGGGAGGCCG GGGCGGGATCT CGCCGGCGCAG  180
GCGGCGGCGG CCGAGGTGGG GTCGCGCGGG GGAGGCGGCT CGAGCTTCGT GCTGCGCGCT   240
CGCTCTTGGG CTCCTCGCTG CAGGAGGGAGT GTGACTATGT GCAGATGATC GAGGTGCAGC   300
ACAAGCAGTG CCTGGAGGAG GCCCAGCTGG AGAATGAGAC AATAGGCTGC AGCAAGATGT   360
GGGACAACCT CACCTGCTGG CCAGCCACCC CTCGGGGCCA GGTAGTTGTC TTGGCCTGTC   420
CCCTCATCTT CAAGCTCTTC TCCTCCATTC AAGGCCGCAA TGTAAGCCGC AGCTGCACCG   480
ACGAAGGCTG GACGCACCTG GAGCCTGGCC CGTACCCCAT TGCCTGTGGT TTGGATGACA   540
AGGCAGCGAG TTTGGATGAG CAGCAGACCA TGTTCTACGG TTCTGTGAAG ACCGGCTACA   600
CCATTGGCTA CGGCCTGTCC CTCGCCACCC TTCTGGTCGC CACAGCTATC CTGAGCCTGT   660
TCAGGAAGCT CCACTGCACG CGGAACTACA TCCACATGCA CCTCTTCATA TCCTTCATCC   720
TGAGGGCTGC CGCTGTCTTC ATCAAAGACT TGGCCCTCTT CGACAGCGGG GAGTCGGACC   780
AGTGCTCCGA GGGCTCGGTG GGCTGTAAGG CAGCCATGGT CTTTTTTCCAA TATTGTGTCA   840
TGGCTAACTT CTTCTGGCTG CTGGTCGAGG GCCTCTACCT GTACACCCTG CTTGCCGTCT   900
CCTTCTTCTC TGAGCGGAAG TACTTCTGGG GGTACATACT CATCGGCTGG GGGGTACCCA   960
GCACATTCAC CATGGTGTGG ACCATCGCCA GGATCCATTT TGAGGATTAT GGTCTGCTCA  1020
GGTGCTGGGA CACCATCAAC TCCTCACTGT GGTGGATCAT AAAGGGCCCC ATCCTCACCT  1080
CCATCTTGGT AAAACTTCATC CTGTTTATTT GCATCATCCG AATCCTGCTT CAGAAACTGC  1140
GGCCCCCAGA TATCAGGAAG AGTGACAGCA GTCCATACTC AAGGCTAGCC AGGTCCACAC  1200
TCCTGCTGAT CCCCCTGTTT GGAGTACACT ACATCATGTT CGCCTTCTTT CCGGACAATT  1260
TTAAGCCTGA AGTGAAGATG GTCTTTGAGC TCGTCGTGGG GTCTTTCCAG GGTTTTGTGG  1320
TGGCTATCCT CTACTGCTTC CTCAATGGTG AGGTGCAGGC GGAGCTGAGG CGGAAGTGGC  1380
GGCGCTGGCA CCTGCAGGGC GTCCTGGGCT GGAACCCCAA ATACCGGCAC CCGTCGGGAG  1440
GCAGCAACGG CGCCACGTGC AGCACGCAGG TTTCCATGCT GACCCGCGTC AGCCCAGGTG  1500
CCCGCCGCTC CTCCAGCTTC CAAGCCGAAG TCTCCCTGGT CTGACCACCA GGATCCCAGC  1560
CCAAGCGGCC CCTCCCGCCC CTTCCCACTC GCAGCAGACG CCGGGACAG AGGCCTGCCC  1620
GGGCGCGCCA GCCCCGGCCC TGGGCTCGGA GGCTGCCCCC GGCCCCCTGG TCTCTGGTCC  1680
GGACACTCCT AGAGAACGCA GCCCTAGAGC CTGCCTGGAG CGTTTCTAGC AAGTGAGAGA  1740
GATGGGAGCT CCTCTCCTGG AGGATGCAGG TGGAACTCAG TCATTAGACT CCTCCTCCAA  1800
AGGCCCCCTA CGCCAATCAA GGGCAAAAAG TCTACATACT TTCATCCTGA CTCTGCCCCC  1860
TGCTGGCTCT TCTGCCCAAT TGGAGGAAAG CAACCGGTGG ATCCTCAAAC AACACTGGTG  1920
TGACCTGAGG GCAGAAAGGT TCTGCCCGGG AAGGTCACCA GCACCAACAC CACGGTAGTG  1980
CCTGAAATTT CACCATTGCT GTCAAGTTCC TTTGGGTTAA GCATTACCAC TCAGGCATTT  2040
GACTGAAGAT GCAGCTCACT ACCCTATTCT CTCTTTACGC TTAGTTATCA GCTTTTTAAA  2100
GTGGGTTATT CTGGAGTTTT TGTTTGGAGA GCACACCTAT CTTAGTGGTT CCCCACCGAA  2160
GTGGACTGGC CCCTGGGTCA GTCTGGTGGG AGGACGGTGC AACCCAAGGA CTGAGGGACT  2220
CTGAAGCCTC TGGGAAATGA GAAGGCAGCC ACCAGCGAAT GCTAGGTCTC GGACTAAGCC  2280
TACCTGCTCT CCAAGTCTCA GTGGCTTCAT CTGTCAAGTG GGACTCCGTC ACACCAGCCA  2340
TTCTTATCTC TCTGTGCTGT GGAAGCAACA GGAATCAAGA GACTGCCCTC CTTGTCCACC  2400
CACCTATGTG CCAACTGTTG TAACTAGGCT CAGAGATGTG CACCCATGGG CTCTGACAGA  2460
AAGCAGATCC TCACCCTGCT ACACATACAG GATTTGAACT CAGATCTGTC TGATAGGAAT  2520
GTGAAAGCAC GGACTCTTAC TGCTAACTTT TGTGTATCGT AACCAGCCAG ATCCTCTTGG  2580
TTATTTGTTT ACCACTTGTA TTATTAATGC CATTATCCCT GAATTCCCCT TGCCACCCCA  2640
CCCTCCCTGG AGTGTGGCTG AGGAGGCCTC CATCTCATGT ATCATCTGGA TAGGAGCCTG  2700
CTGGTCACAG CCTCCTCTGT CTGCCCTTCA CCCCAGTGGC CACTCAGCTT CCTACCCACA  2760
CCTCTGCCAG AAGATCCCCT CAGGACTGCA ACAGGCTTGT GCAACAATAA ATGTTGGCTT  2820
GGAAAAAAAA AAAA
```

75

SEQ ID NO:177 PM72 Protein sequence:
Protein Accession #:   JC2195

```
1          11         21         31         41         51
|          |          |          |          |          |
MPPPPLLSLR RLGGGWSAVT RLVVAAAGAR SRGGRGGSRG AGGGGRGGVA RRRRLELRAA    60
RSLLGSSLQE ECDYVQMIEV QHKQCLEEAQ LENETIGCSK MWDNLTCWPA TPRGQVVVLA   120
```

373

```
CPLIFKLFSS  IQGRNVSRSC  TDEGWTHLEP  GPYPIACGLD  DKAASLDEQQ  TMFYGSVKTG   180
YTIGYGLSLA  TLLVATAILS  LFRKLHCTRN  YIHMHLFISF  ILRAAAVFIK  DLALFDSGES   240
DQCSEGSVGC  KAAMVFFQYC  VMANFFWLLV  EGLYLYTLLA  VSFFSERKYF  WGYILIGWGV   300
PSTFTMVWTI  ARIHFEDYGL  LRCWDTINSS  LWWIIKGPIL  TSILVNFILF  ICIIRILLQK   360
LRPPDIRKSD  SSPYSRLARS  TLLLIPLFGV  HYIMFAFFPD  NFKPEVKMVF  ELVVGSFQGF   420
VVAILYCFLN  GEVQAELRRK  WRRWHLQGVL  GWNPKYRHPS  GGSNGATCST  QVSMLTRVSP   480
GARRSSSFQA  EVSLV
```

SEQ ID NO:178 BFF8 DNA SEQUENCE

Nucleic Acid Accession #:   AL133619
Coding sequence: 1-2070 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGAGCGGTG  CGGGGGTGGC  GGCTGGGACG  CGGCCCCCCA  GCTCGCCGAC  CCCGGGCTCT    60
CGGCGCCGGC  GCCAGCGCCC  CTCTGTGGGC  GTCCAGTCCT  TGAGGCCGCA  GAGCCCGCAG   120
CTCAGGCAGA  GCGACCCGCA  GAAACGGAAC  CTGGACCTGG  AGAAAAGCCT  GCAGTTCCTG   180
CAGCAGCAGC  ACTCGGAGAT  GCTGGCCAAG  CTCCATGAGG  AGATCGAGCA  TCTGAAGCGG   240
GAAAACAAGG  GTGAGCCGGC  GCGGGGCCCT  AGGCCGGCCC  TGCCTCCCCA  GGCACACTCA   300
ACACTGCCGC  TCCCGCAGCA  CAGAAACACA  GCCATCAACT  CCAGCACACG  CCTGGGCTCA   360
GGGGGAACAC  AGGACGGGGA  GCCCCTCCAG  ACTGTCCTTG  CCCACCTGGC  TGCACTGGCC   420
CCTGTATGCC  AACCCAGTGG  GTACAGGTTC  TGGGGGACCT  GGACAGATGC  CGCTACCTCT   480
AGCCGTGGCT  GGACGATGTT  ATGCAGCCAA  GCACAGCACG  TGCTGCTCTC  GGGAAGCCCA   540
GGGCCTGAGG  TCATTGCAGG  GCGGCAGGTG  GCCACAGGGT  GCTCCCCAGA  CCTCCCTCCT   600
CCAAGTAGAG  CTGAAATGGG  AAGGAACCCC  TGGGACAGCC  CCTGCCCTGC  TAGATCTTTG   660
CCTCAGATTG  CTGCTGTGGC  CAGGCCCAGG  ATTTCCAGCC  CTATGGCTCT  GAGTCCTCAC   720
ATGCTGGGGG  CCCAGGGGAT  ATGGACACAC  TCCATCCAGG  GATCCCTTCC  TGCCATCTGG   780
GCAGCAACCA  TGGGGACAAA  GGGAGGAAGC  AGAGTCCTGT  TTCCTTGCCA  CTTGTCCAAG   840
GCACTTCCCC  ATCCTGACAG  CGGCCCCCAC  CCAGCCCAGG  ATCCTGGGCT  GTGGTCTCAA   900
GCTCACTTCC  CATTATCTTT  GGGGCTGGGG  CTGACATCAG  GAGGACATCT  GACTGGTGGA   960
TGGAGCCAGC  CTGGGAACAT  CGCAGCTGGG  GCAGTGCCTA  GGGCTCTCCC  TTCCCAGGGA  1020
GACATGGAGA  AGGGGGTTGA  GGGAGGGCCC  TTCCCTAGCC  GCTGTGGCAA  CTCCAGTGAG  1080
CTGTTCTGGG  CAAAGTGTGG  CCCAAGTCGG  CAGCCCCAGC  CCTGCAGTGC  TGGGGACGCT  1140
GACAGGACAC  GGGAAGAGGC  CATGCTTTCC  CTCGGGACCT  GCTGTTCCAT  GTGTCCCAAG  1200
CCCTCCTGCT  TTCCAGATGG  CCCCTCAGGA  AACCACCTTT  CCAGGGCCTC  TGCTCCCTTG  1260
GGCGCTCGCT  GGGTCTGCAT  CAACGGAGTG  TGGGTAGAGC  CGGGAGGACC  CAGCCCTGCC  1320
AGGCTGAAGG  AGGGCTCCTC  ACGGACACAC  AGGCCAGGAG  GCAAGCGTGG  GCGTCTTGCG  1380
GGCGGTAGCG  CCGACACTGT  GCGCTCTCCT  GCAGACAGCC  TCTCCATGTC  AAGCTTCCAG  1440
TCTGTCAAGT  CCATCTCTAA  TTCAGCCAAC  TCTCAAGGCA  AGGCCAGGCC  CCAGCCCGGC  1500
TCCTTCAACA  AGCAAGATTC  AAAAGCTGAC  GTCTCCCAGA  AGGCGGACCT  GGAAGAGGAG  1560
CCCCTACTTC  ACAACAGCAA  GCTGGACAAA  GTTCCTGGGG  TACAAGGGCA  GGCCAGAAAG  1620
GAGAAAGCAG  AGGCCTCTAA  TGCAGGAGCT  GCCTGTATGG  GGAACAGCCA  GCACCAGGGC  1680
AGGCAGATGG  GGGCGGGGGC  ACACCCCCCA  ATGATCCTGC  CCCTTCCCCT  GCGAAAGCCC  1740
ACCACACTTA  GGCAGTGCGA  AGTGCTCATC  CGCGAGCTGT  GGAATACCAA  CCTCCTGCAG  1800
ACCCAAGAGC  TGCGGCACCT  CAAGTCCCTC  CTGGAAGGGA  GCCAGAGGCC  CCAGGCAGCC  1860
CCGGAGGCAG  CTAGCTTTCC  CAGGGACCAA  GAAGCCACGC  ATTTCCCCAA  GGTCTCCACC  1920
AAGAGCCTCT  CCAAGAAATG  CCTGAGCCCA  CCTGTGGCGG  AGCGTGCCAT  CCTGCCCGCA  1980
CTGAAGCAGA  CCCCGAAGAA  CAACTTTGCC  GAGAGGCAGA  AGAGGCTGCA  GGCAATGCAG  2040
AAACGGCGCC  TGCATCGCTC  AGTGCTTTGA
```

SEQ ID NO:179 BFF8 Protein sequence:
Protein Accession #:   T43457

```
1          11         21         31         41         51
|          |          |          |          |          |
MSGAGVAAGT  RPPSSPTPGS  RRRRQRPSVG  VQSLRPQSPQ  LRQSDPQKRN  LDLEKSLQFL    60
QQQHSEMLAK  LHEEIEHLKR  ENKGEPARGP  RPALPPQAHS  TLPLPQHRNT  AINSSTRLGS   120
GGTQDGEPLQ  TVLAHLAALA  PVCQPSGYRF  WGTWTDAATS  SRGWTMLCSQ  AQHVLLSGSP   180
GPEVIAGRQV  ATGCSPDLPP  PSRAEMGRNP  WDSPCPARSL  PQIAAVARPR  ISSPMALSPH   240
MLGAQGIWTH  SIQGSLPAIW  AATMGTKGGS  RVLFPCHLSK  ALPHPDSGPH  PAQDPGLWSQ   300
AHFPLSLGLG  LTSGGHLTGG  WSQPGNIAAG  AVPRALPSQG  DMEKGVEGGP  FPSRCGNSSE   360
LFWAKCGPSR  QPQPCSAGDA  DRTREEAMLS  LGTCCSMCPK  PSCFPDGPSG  NHLSRASAPL   420
GARWVCINGV  WVEPGGPSPA  RLKEGSSRTH  RPGGKRGRLA  GGSADTVRSP  ADSLSMSSFQ   480
SVKSISNSAN  SQGKARPQPG  SFNKQDSKAD  VSQKADLEEE  PLLHNSKLDK  VPGVQGQARK   540
EKAEASNAGA  ACMGNSQHQG  RQMGAGAHPP  MILPLPLRKP  TTLRQCEVLI  RELWNTNLLQ   600
TQELRHLKSL  LEGSQRPQAA  PEEASFPRDQ  EATHFPKVST  KSLSKKCLSP  PVAERAILPA   660
LKQTPKNNFA  ERQKRLQAMQ  KRRLHRSVL
```

SEQ ID NO:180 BCR4 DNA SEQUENCE

Nucleic Acid Accession #:   NM_012319.2
                          Coding sequence: 138-2405 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
CTCGTGCCGA  ATTCGGCACG  AGACCGCGTG  TTCGCGCCTG  GTAGAGATTT  CTCGAAGACA    60
CCAGTGGGCC  CGTGTGGAAC  CAAACCTGCG  CGCGTGGCCG  GGCCGTGGGA  CAACGAGGCC   120
```

374

```
GCGGAGACGA AGGCGCAATG GCGAGGAAGT TATCTGTAAT CTTGATCCTG ACCTTTGCCC    180
TCTCTGTCAC AAATCCCCTT CATGAACTAA AAGCAGCTGC TTTCCCCCAG ACCACTGAGA    240
AAATTAGTCC GAATTGGGAA TCTGGCATTA ATGTTGACTT GGCAATTTCC ACACGGCAAT    300
ATCATCTACA ACAGCTTTTC TACCGCTATG GAGAAAATAA TTCTTTTGTCA GTTGAAGGGT    360
TCAGAAAATT ACTTCAAAAT ATAGGCATAG ATAAGATTAA AAGAATCCAT ATACACCATG    420
ACCACGACCA TCACTCAGAC CACGAGCATC ACTCAGACCA TGAGCGTCAC TCAGACCATG    480
AGCATCACTC AGACCACGAG CATCACTCTG ACCATGATCA TCACTCTCAC CATAATCATG    540
CTGCTTCTGG TAAAAATAAG CGAAAAGCTC TTTGCCCAGA CCATGACTCA GATAGTTCAG    600
GTAAAGATCC TAGAAACAGC CAGGGGAAAG GAGCTCACCG ACCAGAACAT GCCAGTGGTA    660
GAAGGAATGT CAAGGACAGT GTTAGTGCTA GTGAAGTGAC CTCAACTGTG TACAACACTG    720
TCTCTGAAGG AACTCACTTT CTAGAGACAA TAGAGACTCC AAGACCTGGA AAACTCTTCC    780
CCAAAGATGT AAGCAGCTCC ACTCCACCCA GTGTCACATC AAAGAGCCGG GTGAGCCGGC    840
TGGCTGGTAG GAAAACAAAT GAATCTGTGA GTGAGCCCCG AAAAGGCTTT ATGTATTCCA    900
GAAACACAAA TGAAAATCCT CAGGAGTGTT TCAATGCATC AAAGCTACTG ACATCTCATG    960
GCATGGGCAT CCAGGTTCCG CTGAATGCAA CAGAGTTCAA CTATCTCTGT CCAGCCATCA   1020
TCAACCAAAT TGATGCTAGA TCTTGTCTGA TTCATACAAG TGAAAAGAAG GCTGAAATCC   1080
CTCCAAAGAC CTATTCATTA CAAATAGCCT GGGTTGGTGG TTTTTATAGCC ATTTCCATCA   1140
TCAGTTTCCT GTCTCTGCTG GGGGTTATCT TAGTGCCTCT CATGAATCGG GTGTTTTTCA   1200
AATTTCTCCT GAGTTTCCTT GTGGCACTGG CCGTTGGGAC TTTGAGTGGT GATGCTTTTT   1260
TACACCTTCT TCCACATTCT CATGCAAGTC ACCACCATAG TCATAGCCAT GAAGAACCAG   1320
CAATGGAAAT GAAAAGAGGA CCACTTTTCA GTCATCTGTC TTCTCAAAAC ATAGAAGAAA   1380
GTGCCTATTT TGATTCCACG TGGAAGGGTC TAACAGCTCT AGGAGGCCTG TATTTCATGT   1440
TTCTTGTTGA ACATGTCCTC ACATTGATCA AACAATTTAA AGATAAGAAG AAAAAGAATC   1500
AGAAGAAACC TGAAAATGAT GATGATGTGG AGATTAAGAA GCAGTTGTCC AAGTATGAAT   1560
CTCAACTTTC AACAAATGAG GAGAAAGTAG ATACAGATGA TCGAACTGAA GGCTATTTAC   1620
GAGCAGACTC ACAAGAGCCC TCCCACTTTG ATTCTCAGCA GCCTGCAGTC TTGGAAGAAG   1680
AAGAGGTCAT GATAGCTCAT GCTCATCCAC AGGAAGTCTA CAATGAAATAT GTACCCAGAG   1740
GGTGCAAGAA TAAATGCCAT TCACATTTCC ACGATACACT CGGCCAGTCA GACGATCTCA   1800
TTCACCACCA TCATGACTAC CATCATATTC TCCATCATCA CCACCACCAA AACCACCATC   1860
CTCACAGTCA CAGCCAGCGC TACTCTCGGG AGGAGCTGAA AGATGCCGGC GTCGCCACTT   1920
TGGCCTGGAT GGTGATAATG GGTGATGGCC TGCACAATTT CAGCGATGGC CTAGCAATTG   1980
GTGCTGCTTT TACTGAAGGC TTATCAAGTG GTTTAAGTAC TTCTGTTGCT GTGTTCTGTC   2040
ATGAGTTGCC TCATGAATTA GGTGACTTTG CTGTTCTACT AAAGGCTGGC ATGACCGTTA   2100
AGCAGGCTGT CCTTTATAAT GCATTGTCAG CCATGCTGGC GTATCTTGGA ATGGCAACAG   2160
GAATTTTCAT TGGTCATTAT GCTGAAAAATG TTTCTATGTG GATATTTGCA CTTACTGCTG   2220
GCTTATTCAT GTATGTTGCT CTGGTTGATA TGGTACCTGA AATGCTCAC AATGATGCTA   2280
GTGACCATGG ATGTAGCCGC TGGGGGTATT TCTTTTTACA GAATGCTGGG ATGCTTTTGG   2340
GTTTTGGAAT TATGTTACTT ATTTCCATAT TTGAACATAA AATCGTGTTT CGTATAAATT   2400
TCTAGTTAAG GTTTAAATGC TAGAGTAGCT TAAAAGTTG TCATAGTTTC AGTAGGTCAT   2460
AGGGAGATGA GTTTGTATGC TGTACTATGC AGCGTTTAAA GTTAGTGGGT TTTGTGATTT   2520
TTGTATTGAA TATTGCTGTC TGTTACAAAG TCAGTTAAAG GTACGTTTTA ATATTTAAGT   2580
TATTCTATCT TGGAGATAAA ATCTGTATGT GCAATTCACC GGTATTACCA GTTTATTATG   2640
TAAACAAGAG ATTTGGCATG ACATGTTCTG TATGTTTCAG GGAAAAATGT CTTTAATGCT   2700
TTTTCAAGAA CTAACACAGT TATTCCTATA CTGGATTTTA GGTCTCTGAA GAACTGCTGG   2760
TGTTTAGGAA TAAGAATGTG CATGAAGCCT AAAATACCAA GAAAGCTTAT ACTGAATTTA   2820
AGCAAAGAAA TAAAGGAGAA AAGAGAAGAA TCTGAGAATT GGGGAGGCAT AGATTCTTAT   2880
AAAAATCACA AAATTTGTTG TAAATTAGAC GGGAGAAATT TAGAATTAAG TATAAAAAGG   2940
CAGAATTAGT ATAGAGTACA TTCATTAAAC ATTTTTGTCA GGATTATTTC CCGTAAAAAC   3000
GTAGTGAGCA CTCTCATATA CTAATTAGTG TACATTTAAC TTTGTATAAT ACAGAAATCT   3060
AAATATATTT AATGAATTCA AGCAATATAC ACTTGACCAA GAAATTGGAA TTTCAAAATG   3120
TTCGTGCGGG TTATATACCA GATGAGTACA GTGAGTAGTT TATGTATCAC CAGACTGGGT   3180
TATTGCCAAG TTATATATCA CCAAAAGCTG TATGACTGGA TGTTCTGGTT ACCTGGTTTA   3240
CAAAATTATC AGAGTAGTAA AACTTTGATA TATATGAGGA TATTAAAACT ACACTAAGTA   3300
TCATTTGATT CGATTCAGAA AGTACTTTGA TATCTCTCAG TGCTTCAGTG CTATCATTGT   3360
GAGCAATTGT CTTTATATAC GGTACTGTAG CCATACTAGG CCTGTCTGTG GCATTCTCTA   3420
GATGTTTCTT TTTTACACAA TAAATTCCTT ATATCAGCTT G
```

```
   1         11         21         31         41         51
   |         |          |          |          |          |
MARKLSVILI LTFALSVTNP LHELKAAAFP QTTEKISPNW ESGINVDLAI STRQYHLQQL    60
FYRYGENNSL SVEGFRKLLQ NIGIDKIKRI HIHHDHDHHS DHEHHSDHER HSDHEHHSDH   120
EHHSDHDHHS HHNHAASGKN KRKALCPDHD SDSSGKDPRN SQGKGAHRPE HASGRRNVKD   180
SVSASEVTST VYNTVSEGTH FLETIETPRP GKLFPKDVSS STPPSVTSKS RVSRLAGRKT   240
NESVSEPRKG FMYSRNTNEN PQECFNASKL LTSHGMGIQV PLNATEFNYL CPAIINQIDA   300
RSCLIHTSEK KAEIPPKTYS LQIAWVGGFI AISIISFLSL LGVILVPLMN RVFFKFLLSF   360
LVALAVGTLS GDAFLHLLPH SHASHHHSHS HEEPAMEMKR GPLFSHLSSQ NIEESAYFDS   420
TWKGLTALGG LYFMFLVEHV LTLIKQFKDK KKKNQKKPEN DDDVEIKKQL SKYESQLSTN   480
EEKVDTDDRT EGYLRADSQE PSHFDSQQPA VLEEEEVMIA HAHPQEVYNE YVPRGCKNKC   540
HSHFHDTLGQ SDDLIHHHHD YHHILHHHHH QNHHPHSHSQ RYSREELKDA GVATLAWMVI   600
MGDGLHNFSD GLAIGAAFTE GLSSGLSTSV AVFCHELPHE LGDFAVLLKA GMTVKQAVLY   660
NALSAMLAYL GMATGIFIGH YAENVSMWIF ALTAGLFMYV ALVDMVPEML HNDASDHGCS   720
RWGYFFLQNA GMLLGFGIML LISIFEHKIV FRINF
```

375

SEQ ID NO:182 BCY2 DNA sequence
Nucleic Acid Accession #:   NM_001203
Coding sequence: 274-1782 (underlined sequences correspond to start and stop codons)

```
1     11    21    31    41    51
|     |     |     |     |     |
CGCGGGGCGC GGAGTCGGCG GGGCCTCGCG GGACGCGGGC AGTGCGGAGA CCGCGGCGCT  60
GAGGACGCGG GAGCCGGGAG CGCACGCGCG GGGTGGAGTT CAGCCTACTC TTTCTTAGAT  120
GTGAAAGGAA AGGAAGATCA TTTCATGCCT TGTTGATAAA GGTTCAGACT TCTGCTGATT   180
CATAACCATT TGGCTCTGAG CTATGACAAG AGAGGAAACA AAAAGTTAAA CTTACAAGCC  240
TGCCATAAGT GAGAAGCAAA CTTCCTTGAT AACATGCTTT TGCGAAGTGC AGGAAAATTA  300
AATGTGGGCA CCAAGAAAGA GGATGGTGAG AGTACAGCCC CCACCCCCCG TCCAAAGGTC  360
TTGCGTTGTA AATGCCACCA CCATTGTCCA GAAGACTCAG TCAACAATAT TTGCAGCACA  420
GACGGATATT GTTTCACGAT GATAGAAGAG GATGACTCTG GGTTGCCTGT GGTCACTTCT  480
GGTTGCCTAG GACTAGAAGG CTCAGATTTT CAGTGTCGGG ACACTCCCAT TCCTCATCAA  540
AGAAGATCAA TTGAATGCTG CACAGAAAGG AACGAATGTA ATAAAGACCT ACACCCTACA  600
CTGCCTCCAT TGAAAAACAG AGATTTTGTT GATGGACCTA TACACCACAG GGCTTTACTT  660
ATATCTGTGA CTGTCTGTAG TTTGCTCTTG GTCCTTATCA TATTATTTTG TTACTTCCGG  720
TATAAAAGAC AAGAAACCAG ACCTCGATAC AGCATTGGGT TAGAACAGGA TGAAACTTAC  780
ATTCCTCCTG GAGAATCCCT GAGAGACTTA ATTGAGCAGT CTCAGAGCTC AGGAAGTGGA  840
TCAGGCCTCC CTCTGCTGGT CCAAAGGACT ATAGCTAAGC AGATTCAGAT GGTGAAACAG  900
ATTGGAAAAG GTCGCTATGG GGAAGTTTGG ATGGGAAAGT GGCGTGGCGA AAAGGTAGCT    960
GTGAAAGTGT TCTTCACCAC AGAGGAAGCC AGCTGGTTCA GAGAGACAGA AATATATCAG  1020
ACAGTGTTGA TGAGGCATGA AAACATTTTG GGTTTCATTG CTGCAGATAT CAAAGGGACA  1080
GGGTCCTGGA CCCAGTTGTA CCTAATCACA GACTATCATG AAAATGGTTC CCTTTATGAT  1140
TATCTGAAGT CCACCACCCT AGACGCTAAA TCAATGCTGA AGTTAGCCTA CTCTTCTGTC  1200
AGTGGCTTAT GTCATTTACA CACAGAAATC TTTAGTACTC AAGGCAAACC AGCAATTGCC  1260
CATCGAGATC TGAAAAGTAA AAACATTCTG GTGAAGAAAA ATGGAACTTG CTGTATTGCT  1320
GACCTGGGCC TGGCTGTTAA ATTTATTAGT GATACAAATG AAGTTGACAT ACCACCTAAC  1380
ACTCGAGTTG GCACCAAACG CTATATGCCT CCAGAAGTGT TGGACGAGAG CTTGAACAGA  1440
AATCACTTCC AGTCTTACAT CATGGCTGAC ATGTATAGTT TTGGCCTCAT CCTTTGGGAG  1500
GTTGCTAGGA GATGTGTATC AGGAGGTATA GTGGAAGAAT ACCAGCTTCC TTATCATGAC  1560
CTAGTGCCCA GTGACCCCTC TTATGAGGAC ATGAGGGAGA TTGTGTGCAT CAAGAAGTTA  1620
CGCCCCTCAT TCCCAAACCG GTGGAGCAGT GATGAGTGTC TAAGGCAGAT GGGGAAACTC  1680
ATGACAGAAT GCTGGGCTCA CAATCCTGCA TCAAGGCTGA CAGCCCTGCG GGTTAAGAAA  1740
ACACTTGCCA AAATGTCAGA GTCCCAGGAC ATTAAACTCT GATAGGAGAG GAAAAGTAAG  1800
CATCTCTGCA GAAAGCCAAC AGGTACTCTT CTGTTTGTGG GCAGAGCAAA AGACATCAAA  1860
TAAGCATCCA CAGTACAAGC CTTGAACATC GTCCTGCTTC CCAGTGGGTT CAGACCTCAC  1920
CTTTCAGGGA GCGACCTGGG CAAAGACAGA GAAGCTCCCA GAAGGAGAGA TTGATCCGTG  1980
TCTGTTTGTA GGCGGAGAAA CCGTTGGGTA ACTTGTTCAA GATATGATGC AT
```

SEQ ID NO:183 BCY2 Protein sequence
Protein Accession #:   NP_001194

```
1     11    21    31    41    51
|     |     |     |     |     |
MLLRSAGKLN VGTKKEDGES TAPTPRPKVL RCKCHHHCPE DSVNNICSTD GYCFTMIEED   60
DSGLPVVTSG CLGLEGSDFQ CRDTPIPHQR RSIECCTERN ECNKDLHPTL PPLKNRDFVD  120
GPIHHRALLI SVTVCSLLLV LIILFCYFRY KRQETRPRYS IGLEQDETYI PPGESLRDLI  180
EQSQSSGSGS GLPLLVQRTI AKQIQMVKQI GKGRYGEVWM GKWRGEKVAV KVFFTTEEAS  240
WFRETEIYQT VLMRHENILG FIAADIKGTG SWTQLYLITD YHENGSLYDY LKSTTLDAKS  300
MLKLAYSSVS GLCHLHTEIF STQGKPAIAH RDLKSKNILV KKNGTCCIAD LGLAVKFISD  360
TNEVDIPPNT RVGTKRYMPP EVLDESLNRN HFQSYIMADM YSFGLILWEV ARRCVSGGIV  420
EEYQLPYHDL VPSDPSYEDM REIVCIKKLR PSFPNRWSSD ECLRQMGKLM TECWAHNPAS  480
RLTALRVKKT LAKMSESQDI KL
```

SEQ ID NO:184 CBF9 DNA sequence

Nucleic Acid Accession #:   AC005383
Coding Sequence: 328-2751 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
GACAGTGTTC GCGGCTGCAC CGCTCGGAGG CTGGGTGACC CGCGTAGAAG TGAAGTACTT   60
TTTTATTTGC AGACCTGGGC CGATGCCGCT TTAAAAAACG CGAGGGGCTC TATGCACCTC  120
CCTGGCGGTA GTTCCTCCGA CCTCAGCCGG GTCGGGTCGT GCCGCCCTCT CCCAGGAGAG  180
ACAAACAGGT GTCCCACGTG GCAGCCGCGC CCCGGGCGCC CCTCCTGTGA TCCCGTAGCG  240
CCCCCTGGCC CGAGCCGCGC CCGGGTCTGT GAGTAGAGCC GCCCGGGCAC CGAGCGCTGG  300
TCGCCGCTCT CCTTCCGTTA TATCAACATG CCCCCTTTCC TGTTGCTGGA GGCCGTCTGT  360
GTTTTCCTGT TTTCCAGAGT GCCCCCATCT CTCCCTCTCC AGGAAGTCCA TGTAAGCAAA  420
GAAACCATCG GGAAGATTTC AGCTGCCAGC AAAATGATGT GGTGCTCGGC TGCAGTGGAC  480
ATCATGTTTC TGTTAGATGG GTCTAACAGC GTCGGGAAAG GGAGCTTTGA AAGGTCCAAG  540
CACTTTGCCA TCACAGTCTG TGACGGTCTG GACATCAGCC CCGAGAGGGT CAGAGTGGGA  600
GCATTCCAGT TCAGTTCCAC TCCTCATCTG GAATTCCCCT TGGATTCATT TTCAACCCAA  660
```

376

```
      CAGGAAGTGA AGGCAAGAAT CAAGAGGATG GTTTTCAAAG GAGGGCGCAC GGAGACGGAA   720
      CTTGCTCTGA AATACCTTCT GCACAGAGGG TTGCCTGGAG GCAGAAATGC TTCTGTGCCC   780
      CAGATCCTCA TCATCGTCAC TGATGGGAAG TCCCAGGGGG ATGTGGCACT GCCATCCAAG   840
      CAGCTGAAGG AAAGGGGTGT CACTGTGTTT GCTGTGGGGG TCAGGTTTCC CAGGTGGGAG   900
5     GAGCTGCATG CACTGGCCAG CGAGCCTAGA GGGCAGCACG TGCTGTTGGC TGAGCAGGTG   960
      GAGGATGCCA CCAACGGCCT CTTCAGCACC CTCAGCAGCT CGGCCATCTG CTCCAGCGCC  1020
      ACGCCAGACT GCAGGGTCGA GGCTCACCCC TGTGAGCACA GGACGCTGGA GATGGTCCGG  1080
      GAGTTCGCTG GCAATGCCCC ATGCTGGAGA GGATCGCGGC GGACCCTTGC GGTGCTGGCT  1140
      GCACACTGTC CCTTCTACAG CTGGAAGAGA GTGTTCCTAA CCCACCCTGC CACCTGCTAC  1200
10    AGGACCACCT GCCCAGGCCC CTGTGACTCG CAGCCCTGCC AGAATGGAGG CACATGTGTT  1260
      CCAGAAGGAC TGGACGGCTA CCAGTGCCTC TGCCCGCTGG CCTTTGGAGG GGAGGCTAAC  1320
      TGTGCCCTGA AGCTGAGCCT GGAATGCAGG GTCGACCTCC TCTTCCTGCT GGACAGCTCT  1380
      GCGGGCACCA CTCTGGACGG CTTCCTGCGG GCCAAAGTCT TCGTGAAGCG GTTTGTGCGG  1440
      GCCGTGCTGA GCGAGGACTC TCGGGCCCGA GTGGGTGTGG CCACATACAG CAGGGAGCTG  1500
15    CTGGTGGCGG TGCCTGTGGG GGAGTACCAG GATGTGCCTG ACCTGGTCTG GAGCCTCGAT  1560
      GGCATTCCCT TCCGTGGTGG CCCCACCCTG ACGGGCAGTG CCTTGCGGCA GGCGGCAGAG  1620
      CGTGGCTTCG GGAGCGCCAC CAGGACAGGC CAGGACCGGC CACGTAGAGT GGTGGTTTTG  1680
      CTCACTGAGT CACACTCCGA GGATGAGGTT GCGGGCCCAG CGCGTCACGC AAGGGCGCGA  1740
      GAGCTGCTCC TGCTGGGTGT AGGCAGTGAG GCCGTGCGGG CAGAGCTGGA GGAGATCACA  1800
20    GGCAGCCCAA AGCATGTGAT GGTCTACTCG GATCCTCAGG ATCTGTTCAA CCAAATCCCT  1860
      GAGCTGCAGG GGAAGCTGTG CAGCCGGCAG CGGCCAGGGT GCCGGACACA AGCCCTGGAC  1920
      CTCGTCTTCA TGTTGGACAC CTCTGCCTCA GTAGGGCCCG AGAATTTTGC TCAGATGCAG  1980
      AGCTTTGTGA GAAGCTGTGC CCTCCAGTTT GAGGTGAACC CTGACGTGAC ACAGGTCGGC  2040
25    CTGGTGGTGT ATGGCAGCCA GGTGCAGACT GCCTTCGGGC TGGACACCAA ACCCACCCGG  2100
      GCTGCGATGC TGCGGGCCAT TAGCCAGGCC CCCTACCTAG GTGGGGTGGG CTCAGCCGGC  2160
      ACCGCCCTGC TGCACATCTA TGACAAAGTG ATGACCGTCC AGAGGGGTGC CCGGCCTGGT  2220
      GTCCCCAAGG CTGTGGTGGT GCTCACAGGC GGGAGAGGCG CAGAGGATGC AGCCGTTCCT  2280
      GCCCAGAAGC TGAGGAACAA TGGCATCTCT GTCTTGGTCG TGGGCGTGGG GCCTGTCCTA  2340
30    AGTGAGGGTC TGCGGAGGCT TGCAGGTCCC CGGGATTCCC TGATCCACGT GGCAGCTTAC  2400
      GCCGACCTGC GGTACCACCA GGACGTGCTC ATTGAGTGGC TGTGTGGAGA AGCCAAGCAG  2460
      CCAGTCAACC TCTGCAAACC CAGCCCGTGC ATGAATGAGG GCAGCTGCGT CCTGCAGAAT  2520
      GGGAGCTACC GCTGCAAGTG TCGGGATGGC TGGGAGGGCC CCCACTGCGA GAACCGTGAG  2580
      TGGAGCTCTT GCTCTGTATG TGTGAGCCAG GGATGGATTC TTGAGACGCC CCTGAGGCAC  2640
35    ATGGCTCCCG TGCAGGAGGG CAGCAGCCGT ACCCCTCCCA GCAACTACAG AGAAGGCCTT  2700
      GGCACTGAAA TGGTGCCTAC CTTCTGGAAT GTCTGTGCCC CAGGTCCTTA GAATGTCTGC  2760
      TTCCCGCCGT GGCCAGGACC ACTATTCTCA CTGAGGGAGG AGGATGTCCC AACTGCAGCC  2820
      ATGCTGCTTA GAGACAAGAA AGCAGCTGAT GTCACCCACA AACGATGTTG TTGAAAAGTT  2880
      TTGATGTGTA AGTAAATACC CACTTTCTGT ACCTGCTGTG CCTTGTTGAG GCTATGTCAT  2940
40    CTGCCACCTT TCCCTTGAGG ATAAACAAGG GGTCCTGAAG ACTTAAATTT AGCGGCCTGA  3000
      CGTTCCTTTG CACACAATCA ATGCTCGCCA GAATGTTGTT GACACAGTAA TGCCCAGCAG  3060
      AGGCCTTTAC TAGAGCATCC TTTGGACGGC GAAGGCCACG GCCTTTCAAG ATGGAAAGCA  3120
      GCAGCTTTTC CACTTCCCCA GAGACATTCT GGATGCATTT GCATTGAGTC TGAAAGGGGG  3180
45    CTTGAGGGAC GTTTGTGACT TCTTGGCGAC TGCCTTTTGT GTGTGGAAGA GACTTGGAAA  3240
      GGTCTCAGAC TGAATGTGAC CAATTAACCA GCTTGGTTGA TGATGGGGGA GGGGCTGAGT  3300
      TGTGCATGGG CCCAGGTCTG GAGGGCCACG TAAAATCGTT CTGAGTCGTG AGCAGTGTCC  3360
      ACCTTGAAGG TCTTC
```

```
        1          11         21         31         41         51
        |          |          |          |          |          |
      MPPFLLLEAV CVFLFSRVPP SLPLQEVHVS KETIGKISAA SKMMWCSAAV DIMFLLDGSN   60
      SVGKGSFERS KHFAITVCDG LDISPERVRV GAFQFSSTPH LEFPLDSFST QQEVKARIKR  120
      MVFKGGRTET ELALKYLLHR GLPGGRNASV PQILIIVTDG KSQGDVALPS KQLKERGVTV  180
      FAVGVRFPRW EELHALASEP RGQHVLLAEQ VEDATNGLFS TLSSSAICSS ATPDCRVEAH  240
      PCEHRTLEMV REFAGNAPCW RGSRRTLAVL AAHCPFYSWK RVFLTHPATC YRTTCPGPCD  300
      SQPCQNGGTC VPEGLDGYQC LCPLAFGGEA NCALKLSLEC RVDLLFLLDS SAGTTLDGFL  360
      RAKVFVKRFV RAVLSEDSRA RVGVATYSRE LLVAVPVGEY QDVPDLVWSL DGIPFRGGPT  420
      LTGSALRQAA ERGFGSATRT GQDRPRRVVV LLTESHSEDE VAGPARHARA RELLLLGVGS  480
      EAVRAELEEI TGSPKHVMVY SDPQDLFNQI PELQGKLCSR QRPGCRTQAL DLVFMLDTSA  540
      SVGPENFAQM QSFVRSCALQ FEVNPDVTQV GLVVYGSQVQ TAFGLDTKPT RAAMLRAISQ  600
      APYLGGVGSA GTALLHIYDK VMTVQRGARP GVPKAVVVLT GGRGAEDAAV PAQKLRNNGI  660
      SVLVVGVGPV LSEGLRRLAG PRDSLIHVAA YADLRYHQDV LIEWLCGEAK QPVNLCKPSP  720
      CMNEGSCVLQ NGSYRCKCRD GWEGPHCENR EWSSCSVCVS QGWILETPLR HMAPVQEGSS  780
      RTPPSNYREG LGTEMVPTFW NVCAPGP
```

SEQ ID NO:186 PAV1 DNA sequence

Nucleic Acid Accession #:   AF272890
Coding Sequence: 87-1520 (underlined sequences correspond to start and stop codons)

```
        1          11         21         31         41         51
        |          |          |          |          |          |
      TGCTACCCGC GCCCGGGCTT CTGGGGTGTT CCCCAACCAC GGCCCAGCCC TGCCACACCC   60
      CCCGCCCCCG GCCTCCGCAG CTCGGCATGG GCGCGGGGGT GCTCGTCCTG GGCGCCTCCG  120
      AGCCCGGTAA CCTGTCGTCG GCCGCACCGC TCCCCGACGG CGCGGCCACC GCGGCGCGGC  180
```

377

```
TGCTGGTGCC CGCGTCGCCG CCCGCCTCGT TGCTGCCTCC CGCCAGCGAA AGCCCCGAGC    240
CGCTGTCTCA GCAGTGGACA GCGGGCATGG GTCTGCTGAT GGCGCTCATC GTGCTGCTCA    300
TCGTGGCGGG CAATGTGCTG GTGATCGTGG CCATCGCCAA GACGCCGCGG CTGCAGACGC    360
TCACCAACCT CTTCATCATG TCCCTGGCCA GCGCCGACCT GGTCATGGGG CTGCTGGTGG    420
TGCCGTTCGG GGCCACCATC GTGGTGTGGG GCCGCTGGGA GTACGGCTCC TTCTTCTGCG    480
AGCTGTGGAC CTCAGTGGAC GTGCTGTGCG TGACGGCCAG CATCGAGACC CTGTGTGTCA    540
TTGCCCTGGA CCGCTACCTC GCCATCACCT CGCCCTTCCG CTACCAGAGC CTGCTGACGC    600
GCGCGCGGGC GCGGGGCCTC GTGTGCACCG TGTGGGCCAT CTCGGCCCTG GTGTCCTTCC    660
TGCCCATCCT CATGCACTGG TGGCGGGCGG AGAGCGACGA GGCGCGCCGC TGCTACAACG    720
ACCCCAAGTG CTGCGACTTC GTCACCAACC GGGCCTACGC CATCGCCTCG TCCGTAGTCT    780
CCTTCTACGT GCCCCTGTGC ATCATGGCCT TCGTGTACCT GCGGGTGTTC CGCGAGGCCC    840
AGAAGCAGGT GAAGAAGATC GACAGCTGCG AGCGCCGTTT CCTCGGCGGC CCAGCGCGGC    900
CGCCCTCGCC CTCGCCCTCG CCCGTCCCCG CGCCCGCGCC GCCGCCCGGA CCCCCGCGCC    960
CCGCCGCCGC CGCCGCCACC GCCCCGCTGG CCAACGGGCG TGCGGGTAAG CGGCGGCCCT   1020
CGCGCCTCGT GGCCCTACGC GAGCAGAAGG CGCTCAAGAC GCTGGGCATC ATCATGGGCG   1080
TCTTCACGCT CTGCTGGCTG CCCTTCTTCC TGGCCAACGT GGTGAAGGCC TTCCACCGCG   1140
AGCTGGTGCC CGACCGCCTC TTCGTCTTCT TCAACTGGCT GGGCTACGCC AACTCGGCCT   1200
TCAACCCCAT CATCTACTGC CGCAGCCCCG ACTTCCGCAA GGCCTTCCAG GGACTGCTCT   1260
GCTGCGCGCG CAGGGCTGCC CGCCGGCGCC ACGCGACCCA CGGAGACCGG CCGCGCGCCT   1320
CGGGCTGTCT GGCCCGGCCC GGACCCCCGC CATCGCCCGG GGCCGCCTCG GACGACGACG   1380
ACGACGATGT CGTCGGGGCC ACGCCGCCCG CGCGCCTGCT GGAGCCCTGG GCCGGCTGCA   1440
ACGGCGGGGC GGCGGCGGAC AGCGACTCGA GCCTGGACGA GCCGTGCCGC CCCGGCTTCG   1500
CCTCGGAATC CAAGGTGTAG GGCCCGGCGC GGGGCGCGGA CTCCGGGCAC GGCTTCCCAG   1560
GGGAACGAGG AGATCTGTGT TTACTTAAGA CCGATAGCAG GTGAACTCGA AGCCCACAAT   1620
CCTCGTCTGA ATCATCCGAG GCAAAGAGAA AAGCCACGGA CCGTTGCACA AAAAGGAAAG   1680
TTTGGGAAGG GATGGGGAGAG TGGCTTGCTG ATGTTCCTTG TTG
```

```
1          11         21         31         41         51
|          |          |          |          |          |
MGAGVLVLGA SEPGNLSSAA PLPDGAATAA RLLVPASPPA SLLPPASESP EPLSQQWTAG     60
MGLLMALIVL LIVAGNVLVI VAIAKTPRLQ TLTNLFIMSL ASADLVMGLL VVPFGATIVV    120
WGRWEYGSFF CELWTSVDVL CVTASIETLC VIALDRYLAI TSPFRYQSLL TRARARGLVC    180
TVWAISALVS FLPILMHWWR AESDEARRCY NDPKCCDFVT NRAYAIASSV VSFYVPLCIM    240
AFVYLRVFRE AQKQVKKIDS CERRFLGGPA RPPSPSPSPV PAPAPPPGPP RPAAAAATAP    300
LANGRAGKRR PSRLVALREQ KALKTLGIIM GVFTLCWLPF FLANVVKAFH RELVPDRLFV    360
FFNWLGYANS AFNPIIYCRS PDFRKAFQGL LCCARRAARR RHATHGDRPR ASGCLARPGP    420
PPSPGAASDD DDDDVVGATP PARLLEPWAG CNGGAAADSD SSLDEPCRPG FASESKV
```

```
1    11   21   31   41   51
|    |    |    |    |    |
GGCGTCCGCG CACACCTCCC CGCGCCGCCG CCGCCACCGC CCGCACTCCG CCGCCTCTGC    60
CCGCAACCGC TGAGCCATCC ATGGGGGTCG CGGGCCGCAA CCGTCCCGGG GCGGCCTGGG   120
CGGTGCTGCT GCTGCTGCTG CTGCTGCCGC CACTGCTGCT GCTGGCGGGG GCCGTCCCGC   180
CGGGTCGGGG CCGTGCCGCG GGGCCGCAGG AGGATGTAGA TGAGTGTGCC CAAGGGCTAG   240
ATGACTGCCA TGCCGACGCC CTGTGTCAGA ACACACCCAC CTCCTACAAG TGCTCCTGCA   300
AGCCTGGCTA CCAAGGGGAA GGCAGGCAGT GTGAGGACAT CGATGAATGT GGAAATGAGC   360
TCAATGGAGG CTGTGTCCAT GACTGTTTGA ATATTCCAGG CAATTATCGT TGCACTTGTT   420
TTGATGGCTT CATGTTGGCT CATGACGGTC ATAATTGTGT TGATGTGGAC GGTGTGCCTGG   480
AGAACAATGG CGGCTGCCAG CATACCTGTG TCAACGTCAT GGGGAGCTAT GAGTGCTGCT   540
GCAAGGAGGG GTTTTTCCTG AGTGACAATC AGCACACCTG CATTCACCGC TCGGAAGAGG   600
GCCTGAGCTG CATGAATAAG GATCACGGCT GTAGTCACAT CTGCAAGGAG GCCCCAAGGG   660
GCAGCGTCGC CTGTGAGTGC AGGCCTGGTT TTGAGCTGGC CAAGAACCAG AGAGACTGCA   720
TCTTGACCTG TAACCATGGG AACGGTGGGT GCCAGCACTC CTGTGACGAT ACAGCCGATG   780
GCCCAGAGTG CAGCTGCCAT CCACAGTACA AGATGCACAC AGATGGGAGG AGCTGCCTTG   840
AGCGAGAGGA CACTGTCCTG GAGGTGACAG AGAGCAAACAC CACATCAGTG GTGGATGGGG   900
ATAAACGGGT GAAACGGCGG CTGCTCATGG AAACGTGTGC TGTCAACAAT GGAGGCTGTG   960
ACCGCACCTG TAAGGATACT TCGACAGGTG TCCACTGCAG TTGTCCTGTT GGATTCACTC  1020
TCCAGTTGGA TGGGAAGACA TGTAAAGATA TTGATGAGTG CCAGACCCGC AATGGAGGTT  1080
GTGATCATTT CTGCAAAAAC ATCGTGGGCA GTTTTGACTG CGGCTGCAAG AAAGGATTTA  1140
AATTATTAAC AGATGAGAAG TCTTGCCAAG ATGTGGATGA GTGCTCTTTG GATAGGACCT  1200
GTGACCACAG CTGCATCAAC CACCCTGGCA CATTTGCTTG TGCTTGCAAC CGAGGGTACA  1260
CCCTGTATGG CTTCACCCAC TGTGGAGACA CCAATGAGTG CAGCATCAAC AACGGAGGCT  1320
GTCAGCAGGT CTGTGTGAAC ACAGTGGGCA GCTATGAATG CCAGTGCCAC CCTGGGTACA  1380
AGCTCCACTG GAATAAAAAA GACTGTGTGG AAGTGAAGGG GCTCCTGCCC ACAAGTGTGT  1440
CACCCCGTGT GTCCCTGCAC TGCGGTAAGA GTGGTGGAGG AGACGGGTGC TTCCTCAGAT  1500
GTCACTCTGG CATTCACCTC TCTTCAGATG TCACCACCAT CAGGACAAGT GTAACCTTTA  1560
AGCTAAATGA AGGCAAGTGT AGTTTGAAAA ATGCTGAGCT GTTTCCCGAG GGTCTGCGAC  1620
CAGCACTACC AGAGAAGCAC AGCTCAGTAA AAGAGAGCTT CCGCTACGTA AACCTTACAT  1680
GCAGCTCTGG CAAGCAAGTC CCAGGAGCCC CTGGCCGACC AAGCACCCCT AAGGAAATGT  1740
TTATCACTGT TGAGTTTGAG CTTGAAACTA ACCAAAAGGA GGTGACAGCT TCTTGTGACC  1800
```

378

TGAGCTGCAT CGTAAAGCGA ACCGAGAAGC GGCTCCGTAA AGCCATCCGC ACGCTCAGAA 1860
AGGCCGTCCA CAGGGAGCAG TTTCACCTCC AGCTCTCAGG CATGAACCTC GACGTGGCTA 1920
AAAAGCCTCC CAGAACATCT GAACGCCAGG CAGAGTCCTG TGGAGTGGGC CAGGGTCATG 1980
CAGAAAACCA ATGTGTCAGT TGCAGGGCTG GGACCTATTA TGATGGAGCA CGAGAACGCT 2040
GCATTTTATG TCCAAATGGA ACCTTCCAAA ATGAGGAAGG ACAAATGACT TGTGAACCAT 2100
GCCCAAGACC AGGAAATTCT GGGGCCCTGA AGACCCCAGA AGCTTGGAAT ATGTCTGAAT 2160
GTGGAGGTCT GTGTCAACCT GGTGAATATT CTGCAGATGG CTTTGCACCT TGCCAGCTCT 2220
GTGCCCTGGG CACGTTCCAG CCTGAAGCTG GTCGAACTTC CTGCTTCCCC TGTGGAGGAG 2280
GCCTTGCCAC CAAACATCAG GGAGCTACTT CCTTTCAGGA CTGTGAAACC AGAGTTCAAT 2340
GTTCACCTGG ACATTTCTAC AACACCACCA CTCACCGATG TATTCGTTGC CCAGTGGGAA 2400
CATACCAGCC TGAATTTGGA AAAAATAATT GTGTTTCTTG CCCAGGAAAT ACTACGACTG 2460
ACTTTGATGG CTCCACAAAC ATAACCCAGT GTAAAAACAG AAGATGTGGA GGGGAGCTGG 2520
GAGATTTCAC TGGGTACATT GAATCCCCAA ACTACCCAGG CAATTACCCA GCCAACACCG 2580
AGTGTACGTG GACCATCAAC CCACCCCCCA AGCGCCGCAT CCTGATCGTG GTCCCTGAGA 2640
TCTTCCTGCC CATAGAGGAC GACTGTGGGG ACTATCTGGT GATGCGGAAA ACCTCTTCAT 2700
CCAATTCTGT GACAACATAT GAAACCTGCC AGACCTACGA ACGCCCCATC GCCTTCACCT 2760
CCAGGTCAAA GAAGCTGTGG ATTCAGTTCA AGTCCAATGA AGGGAACAGC GCTAGAGGGT 2820
TCCAGGTCCC ATACGTGACA TATGATGAGG ACTACCAGGA ACTCATTGAA GACATAGTTC 2880
GAGATGGCAG GCTCTATGCA TCTGAGAACC ATCAGGAAAT ACTTAAGGAT AAGAAACTTA 2940
TCAAGGCTCT GTTTGATGTC CTGGCCCATC CCCAGAACTA TTTCAAGTAC ACAGCCCAGG 3000
AGTCCCGAGA GATGTTTCCA AGATCGTTCA TCCGATTGCT ACGTTCCAAA GTGTCCAGGT 3060
TTTTGAGACC TTACAAATGA CTCAGCCCAC GTGCCACTCA ATACAAATGT TCTGCTATAG 3120
GGTTGGTGGG ACAGAGCTGT CTTCCTTCTG CATGTCAGCA CAGTCGGGTA TTGCTGCCTC 3180
CCGTATCAGT GACTCATTAG AGTTCAATTT TTATAGATAA TACAGATATT TTGGTAAATT 3240
GAACTTGGTT TTTCTTTCCC AGCATCGTGG ATGTAGACTG AGAATGGCTT TGAGTGGCAT 3300
CAGCTTCTCA CTGCTGTGGG CGGATGTCTT GGATAGATCA CGGGCTGGCT GAGCTGGACT 3360
TTGGTCAGCC TAGGTGAGAC TCACCTGTCC TTCTGGGGTC TTACTCCTCC TCAAGGAGTC 3420
TGTAGTGGAA AGGAGGCCAC AGAATAAGCT GCTTATTCTG AAACTTCAGC TTCCTCTAGC 3480
CCGGCCCTCT CTAAGGGAGC CCTCTGCACT CGTGTGCAGG CTCTGACCAG GCAGAACAGG 3540
CAAGAGGGGA GGGAAGGAGA CCCCTGCAGG CTCCCTCCAC CCACCTTGAG ACCTGGGAGG 3600
ACTCAGTTTC TCCACAGCCT TCTCCAGCCT GTGTGATACA AGTTTGATCC CAGGAACTTG 3660
AGTTCTAAGC AGTGCTCGTG AAAAAAAAAA GCAGAAAGAA TTAGAAATAA ATAAAAACTA 3720
AGCACTTCTG GAGACAT

SEQ ID NO:189 BCO2 Protein sequence
Protein Accession #:   CAB92285

```
  1    11    21    31    41    51
  |     |     |     |     |     |
MGVAGRNRPG AAWAVLLLLL LLPPLLLLAG AVPPGRGRAA GPQEDVDECA QGLDDCHADA   60
LCQNTPTSYK CSCKPGYQGE GRQCEDIDEC GNELNGGCVH DCLNIPGNYR CTCFDGFMLA  120
HDGHNCLDVD ECLENNGGCQ HTCVNVMGSY ECCCKEGFFL SDNQHTCIHR SEEGLSCMNK  180
DHGCSHICKE APRGSVACEC RPGFELAKNQ RDCILTCNHG NGGCQHSCDD TADGPECSCH  240
PQYKMHTDGR SCLEREDTVL EVTESNTTSV VDGDKRVKRR LLMETCAVNN GGCDRTCKDT  300
STGVHCSCPV GFTLQLDGKT CKDIDECQTR NGGCDHFCKN IVGSFDCGCK KGFKLLTDEK  360
SCQDVDECSL DRTCDHSCIN HPGTFACACN RGYTLYGFTH CGDTNECSIN NGGCQQVCVN  420
TVGSYECQCH PGYKLHWNKK DCVEVKGLLP TSVSPRVSLH CGKSGGGDGC FLRCHSGIHL  480
SSDVTTIRTS VTFKLNEGKC SLKNAELFPE GLRPALPEKH SSVKESFRYV NLTCSSGKQV  540
PGAPGRPSTP KEMFITVEFE LETNQKEVTA SCDLSCIVKR TEKRLRKAIR TLRKAVHREQ  600
FHLQLSGMNL DVAKKPPRTS ERQAESCGVG QGHAENQCVS CRAGTYYDGA RERCILCPNG  660
TFQNEEGQMT CEPCPRPGNS GALKTPEAWN MSECGGLCQP GEYSADGFAP CQLCALGTFQ  720
PEAGRTSCFP CGGGLATKHQ GATSFQDCET RVQCSPGHFY NTTTHRCIRC PVGTYQPEFG  780
KNNCVSCPGN TTTDFDGSTN ITQCKNRRCG GELGDFTGYI ESPNYPGNYP ANTECTWTIN  840
PPPKRRILIV VPEIFLPIED DCGDYLVMRK TSSSNSVTTY ETCQTYERPI AFTSRSKKLW  900
IQFKSNEGNS ARGFQVPYVT YDEDYQELIE DIVRDGRLYA SENHQEILKD KKLIKALFDV  960
LAHPQNYFKY TAQESREMFP RSFIRLLRSK VSRFLRPYK
```

SEQ ID NO:190 BFG1 DNA sequence
Nucleic Acid Accession #:   AF007170
Coding sequence: 1-1725 (underlined sequences correspond to stop codon)

```
  1    11    21    31    41    51
  |     |     |     |     |     |
AAGGAGGCGG CCTCCGGGAA AAGCGACCGC AGGACTCCTG AGAGCAGCCT CCATGAGGCC   60
CTGGACCAGT GCATGACCGC CCTGGACCTC TTCCTCACCA ACCAGTTCTC AGAAGCACTC  120
AGCTACCTCA AGCCCAGAAC CAAGGAAAGC ATGTACCATC CACTGACATA TGCCACCATC  180
CTGGAGATGC AGGCCATGAT GACCTTTGAC CCTCAGGACA TCCTGCTTGC CGGCAACATG  240
ATGAAGGAGG CACAGATGCT GTGTCAGAGG CACCGGAGGA AGTCTTCTGT AACAGATTCC  300
TTCAGCAGCC TGGTGAACCG CCCCACGCTG GGCCAATTCA CTGAAGAAGA AATCCACGCT  360
GAGGTCTGCT ATGCAGAGTG CCTGCTGCAG CGAGCAGGCC TGACCTTCCT GCAGGACGAG  420
AACATGGTGA GCTTCATCAA AGGCGGCATC AAAGTTCGAA CAGCTACCA GACCTACAAG  480
GAGCTGGACA GCCTTGTTCA GTCCTCACAA TACTGCAAGG GTGAGAACCA CCCGCACTTT  540
GAAGGAGGAG TGAAGCTTGG TGTAGGGGCC TTCAACCTGA CACTGTCCAT GCTTCCTACT  600
AGGATCCTGA GGCTGTTGGA GTTTGTGGGG TTTTCAGGAA ACAAGGACTA TGGGCTGCTG  660
CAGCTGGAGG AGGGAGCGTC AGGGCACAGC TTCCGCTCTG TGCTCTGTGT CATGCTCCTG  720
CTGTGCTACC ACACCTTCCT CACCTTCGTG CTCGGTACTG GGAACGTCAA CATCGAGGAG  780
GCCGAGAAGC TCTTGAAGCC CTACCTGAAC CGGTACCCTA AGGGTGCCAT CTTCCTGTTC  840
TTTGCAGGGA GGATTGAAGT CATTAAAGGC AACATTGATG CAGCCATCCG GCGTTTCGAG  900
```

379

```
        GAGTGCTGTG AGGCCCAGCA GCACTGGAAG CAGTTTCACC ACATGTGCTA CTGGGAGCTG  960
        ATGTGGTGCT TCACCTACAA GGGCCAGTGG AAGATGTCCT ACTTCTACGC CGACCTGCTC 1020
        AGCAAGGAGA ACTGCTGGTC CAAGGCCACC TACATTTACA TGAAGGCCGC CTACCTCAGC 1080
   5    ATGTTTGGGA AGGAGGACCA CAAGCCGTTC GGGGACGACG AAGTGGAATT ATTTCGAGCT 1140
        GTGCCAGGCC TGAAGCTCAA GATTGCTGGG AAATCTCTAC CCACAGAGAA GTTTGCCATC 1200
        CGGAAGTCCC GGCGCTACTT CTCCTCCAAC CCTATCTCGC TGCCAGTGCC TGCTCTGGAA 1260
        ATGATGTACA TCTGGAACGG CTACGCCGTG ATTGGGAAGC AGCCGAAACT CACGGATGGG 1320
        ATACTTGAGA TTATCACTAA GGCTGAAAGA ATGCTGGAGA AAGGCCCAGA GAACGAGTAC 1380
        TCAGTGGATG ACGAGTGCTT GGTGAAATTG TTGAAAGGCC TGTGTCTGAA ATACCTGGGC 1440
  10    CGTGTCCAGG AGGCCGAGGA GAATTTTAGG AGCATCTCTG CCAATGAAAA GAAGATTAAA 1500
        TATGACCACT ACTTGATCCC AAACGCCCTG CTGGAGCTGG CCCTGCTGCT TATGGAGCAA 1560
        GACAGAAACG AAGAGGCCAT CAAACTTTTG GAATCTGCCA AGCAAAACTA CAAGAATTAC 1620
        TCCATGGAGT CAAGGACACA CTTTCGAATC CAGGCAGCCA CACTCCAAGC CAAGTCTTCC 1680
        CTAGAGAACA GCAGCAGATC CATGGTCTCA TCAGTGTCCT TGTAGCTTTG TGCAGCAGTT 1740
  15    CCGGGCTGGA AGACAGAGAC AGCTGGACAG AGCTCCTGAA AACATTTCAA AATACCCCCT 1800
        CCCCCTGCCC TGCCCTGCCT TTGGGGTCCA CCGGCACTCC AGTTGGATGG CACAACATAG 1860
        TGTATCCGTG CAGAAGCCGA GCTGGCATTT TCACCAGTGT AGCCAAGGGC CTTTGCCAAG 1920
        GGCAGAGCAG GTGGAGCCCT CTGCCTGCCC TATCACACAT ACGGGTACTT GCTTTTCACT 1980
        GTGATGTTTA AGAGAATGTA TGAACAGTTT ACATTTTCCT TAGAAATACA TTGATGGGAT 2040
  20    CACAGTTGGC TTTAAAAACC AACAACAATC AACCACCTGT AAGTCTTTGT CTTCACCTAT 2100
        TATCATCTGG AGGTAAATCT CTTTATATGA TGATGCCAAA GGGCAAATTG CTTTTCAAAT 2160
        TCAGCAAGTT CTCAGCTTGT GTGACGGAAG GTCCTTCAGA GGACCTGAGG AATGCCTGGG 2220
        AGAGGCTAAG CCTCAGGCTT CAATGCTTCT GGGGTTGGGC ATGAGGATGT ACACAGACAC 2280
        CCACTACCTT ACTACTCACA CTTCATTTCA CTCCTTTTGT AAATTTCCAA TTTAAAAATC 2340
  25    AAGCACGTCT TTTTAGTGAG ATAAAATCTG AGCTCTTCTG TAGAAAAATC AATCTCTACC 2400
        AGTAGAAAAT GCCAGGGCTT GATGGAAGAG CTGTGTAGCC CTTTCTATGC CAAAGCCAGG 2460
        AAATTTGGGG GGCAGGAGGA GGTTCTCAGA ATCCAGTCTG TATCTTTGCT GTATGCCAAA 2520
        CTGAAACCAC TGGGAATAAT TTATGAAACA TAAAAATCTT CTGTACTTCA CTCCAAGGTA 2580
  30    CATTTGCTTA CTGACAGCAT TTTTGTTAAA ACTGTTATTC TTGAAAAAAA AAAAAAAAAA 2640
        AA
```

```
        1    11    21    31    41    51
        |    |     |     |     |     |
        MTALDLFLTN QFSEALSYLK PRTKESMYHS LTYATILEMQ AMMTFDPQDI LLAGNMMKEA  60
        QMLCQRHRRK SSVTDSFSSL VNRPTLGQFT EEEIHAEVCY AECLLQRAAL TFLQDENMVS 120
        FIKGGIKVRN SYQTYKELDS LVQSSQYCKG ENHPHFEGGV KLGVGAFNLT LSMLPTRILR 180
        LLEFVGFSGN KDYGLLQLEE GASGHSFRSV LCVMLLLCYH TFLTFVLGTG NVNIEEAEKL 240
        LKPYLNRYPK GAIFLFFAGR IEVIKGNIDA AIRRFEECCE AQQHWKQFHH MCYWELMWCF 300
        TYKGQWKMSY FYADLLSKEN CWSKATYIYM KAAYLSMFGK EDHKPFGDDE VELFRAVPGL 360
        KLKIAGKSLP TEKFAIRKSR RYFSSNPISL PVPALEMMYI WNGYAVIGKQ PKLTDGILEI 420
        ITKAEEMLEK GPENEYSVDD ECLVKLLKGL CLKYLGRVQE AEENFRSISA NEKKIKYDHY 480
        LIPNALLELA LLLMEQDRNE EAIKLLESAK QNYKNYSMES RTHFRIQAAT LQAKSSLENS 540
        SRSMVSSVSL
```

```
        1    11    21    31    41    51
        |    |     |     |     |     |
        ATGACTAGGA AGAGGACATA CTGGGTGCCC AACTCTTCTG GTGGCCTCGT GAATCGTGGC  60
        ATCGACATAG GCGATGACAT GGTTTCAGGA CTTATTTATA AAACCTATAC TCTCCAAGAT 120
        GGCCCCTGGA GTCAGCAAGA GAGAAATCCT GAGGCTCCAG GGAGGGCAGC TGTCCCACCG 180
        TGGGGGAAGT ATGATGCTGC CTTGAGAACC ATGATTCCCT TCCGTCCCAA GCCGAGGTTT 240
        CCTGCCCCCC AGCCCCTGGA CAATGCTGGC CTGTTCTCCT ACCTCACCGT GTCATGGCTC 300
        ACCCCGCTCA TGATCCAAAG CTTACGGAGT CGCTTAGATG AGAACACCAT CCCTCCACTG 360
        TCAGTCCATG ATGCCTCAGA CAAAAATGTC CAAAGGCTTC ACCGCCTTTG GGAAGAAGAA 420
        GTCTCAAGGC GAGGGATTGA AAAAGCTTCA GTGCTTCTGG TGATGCTGAG GTTCCAGAGA 480
        ACAAGGTTGA TTTTCGATGC ACTTCTGGGC ATCTGCTTCT GCATTGCCAG TGTACTCGGG      540
        CCAATATTGA TTATACCAAA GATCCTGGAA TATTCAGAAG AGCAGTTGGG GAATGTTGTC 600
        CATGGAGTGG GACTCTGCTT TGCCCTTTTT CTCTCCGAAT GTGTGAAGTC TCTGAGTTTC 660
        TCCTCCAGTT GGATCATCAA CCAACGCACA GCCATCAGGT TCCGAGCAGC TGTTTCCTCC 720
        TTTGCCTTTG AGAAGCTCAT CCAATTTAAG TCTGTAATAC ACATCACCTC AGGAGAGGCC 780
        ATCAGCTTCT TCACCGGTGA TGTAAACTAC CTGTTTGAAG GGGTGTGCTA TGGACCCCTA 840
        GTACTGATCA CCTGCGCATC GCTGGTCATC TGCAGCATTT CTTCCTACTT CATTATTGGA 900
        TACACTGCAT TTATTGCCAT CTTATGCTAT CTCCTGGTTT TCCCACTGGC GGTATTCATG 960
        ACAAGAATGG CTGTGAAGGC TCAGCATCAC ACATCTGAGG TCAGCGACCA GCGCATCCGT 1020
        GTGACCAGTG AAGTTCTCAC TTGCATTAAG CTGATTAAAA TGTACACATG GGAGAAACCA 1080
        TTTGCAAAAA TCATTGAAGG TATGGAAAGT CTGACTTTCT GCTCCAAACC TGGTGATGGC 1140
  75    ATGGCCTTCA GCATGCTGGC CTCCTTGAAT CTCCTTCGGC TGTCAGTGTT CTTTGTGCCT 1200
        ATTGCAGTCA AAGGTCTCAC GAATTCCAAG TCTGCAGTGA TGAGGTTCAA GAAGTTTTTC 1260
        CTCCAGGAGA GCCCTGTTTT CTATGTCCAG ACATTACAAG ACCCCAGCAA AGCTCTGGTC 1320
        TTTGAGGAGG CCACCTTGTC ATGGCAACAG ACCTGTCCCG GGATCGTCAA TGGGGCACTG 1380
        GAGCTGGAGA GGAACGGGCA TGCTTCTGAG GGGATGACCA GGCCTAGAGA TGCCCTCGGG 1440
```

```
CCAGAGGAAG AAGGGAACAG CCTGGGCCCA GAGTTGCACA AGATCAACCT GGTGGTGTCC 1500
AAGGGGATGA TGTTAGGGGT CTGCGGCAAC ACGGGGAGTG GTAAGAGCAG CCTGTTGTCA 1560
GCCATCCTGG AGGAGATGCA CTTGCTCGAG GGCTCGGTGG GGGTGCAGGG AAGCCTGGCC 1620
TATGTCCCCC AGCAGGCCTG GATCGTCAGC GGGAACATCA GGGAGAACAT CCTCATGGGA 1680
GGCGCATATG ACAAGGCCCG ATACCTCCAG GTGCTCCACT GCTGCTCCCT GAATCGGGAC 1740
CTGGAACTTC TGCCCTTTGG AGACATGACA GAGATTGGAG AGCGGGGCCT CAACCTCTCT 1800
GGGGGGCAGA AACAGAGGAT CAGCCTGGCC CGCGTCGTCT ATTCCGACCG TCAGATCTAC 1860
CTGCTGGACG ACCCCCTGTC TGCTGTGGAC GCCCACGTGG GGAAGCACAT TTTTGAGGAG 1920
TGCATTAAGA AGACACTCAG GGGGAAGACG GTCGTCCTGG TGACCCACCA GCTGCAGTAC 1980
TTAGAATTTT GTGGCCAGAT CATTTTGTTG GAAAATGGGA AAATCTGTGA AAATGGAACT 2040
CACAGTGAGT TAATGCAGAA AAAGGGGAAA TATGCCCAAC TTATCCAGAA GATGCACAAG 2100
GAAGCCACTT CGGACATGTT GCAGGACACA GCAAAGATAG CAGAGAAGCC AAAGGTAGAA 2160
AGTCAGGCTC TGGCCACCTC CCTGGAAGAG TCTCTCAACG GAAATGCTGT GCCGGAGCAT 2220
CAGCTCACAC AGGAGGAGGA GATGGAAGAA GGCTCCTTGA GTTGGAGGGT CTACCACCAC 2280
TACATCCAGG CAGCTGGAGG TTACATGGTC TCTTGCATAA TTTTCTTCTT CGTGGTGCTG 2340
ATCGTCTTCT TAACGATCTT CAGCTTCTGG TGGCTGAGCT ACTGGTTGGA GCAGGGCTCG 2400
GGGACCAATA GCAGCCGAGA GAGCAATGGA ACCATGGCAG ACCTGGGCAA CATTGCAGAC 2460
AATCCTCAAC TGTCCTTCTA CCAGCTGGTG TACGGGCTCA ACGCCCTGCT CCTCATCTGT 2520
GTGGGGGTCT GCTCCTCAGG GATTTTCACC AAAGTCACGA GGAAGGCATC CACGGCCCTG 2580
CACAACAAGC TCTTCAACAA GGTTTTCCGC TGCCCCATGA GTTTCTTTGA CACCATCCCA 2640
ATAGGCCGGC TTTTGAACTG CTTCGCAGGG GACTTGGAAC AGCTGGACCA GCTCTTGCCC 2700
ATCTTTTTCAG AGCAGTTCCT GGTCCTGTCC TTAATGGTGA TCGCCGTCCT GTTGATTGTC 2760
AGTGTGCTGT CTCCATATAT CCTGTTAATG GGAGCCATAA TCATGGTTAT TTGCTTCATT 2820
TATTATATGA TGTTCAAGAA GGCCATCGGT GTGTTCAAGA GACTGGAGAA CTATAGCCGG 2880
TCTCCTTTAT TCTCCCACAT CCTCAATTCT CTGCAAGGCC TGAGCTCCAT CCATGTCTAT 2940
GGAAAAACTG AAGACTTCAT CAGCCAGTTT AAGAGGCTGA CTGATGCGCA GAATAACTAC 3000
CTGCTGTTGT TTCTATCTTC CACACGATGG ATGGCATTGA GGCTGGAGAT CATGACCAAC 3060
CTTGTGACCT TGGCTGTTGC CCTGTTCGTG GCTTTTGGCA TTTCCTCCAC CCCCTACTCC 3120
TTTAAAGTCA TGGCTGTCAA CATCGTGCTG CAGCTGGCGT CCAGCTTCCA GGCCACTGCC 3180
CGGATTGGCT TGGAGACAGA GGCACAGTTC ACGGCTGTAG AGAGGATACT GCAGTACATG 3240
AAGATGTGTG TCTCGGAAGC TCCTTTACAC ATGGAAGGCA CAAGTTGTCC CCAGGGGTGG 3300
CCACAGCATG GGGAAATCAT ATTTCAGGAT TATCACATGA AATACAGAGA CAACACACCC 3360
ACCGTGCTTC ACGGCATCAA CCTGACCATC CGCGGCCACG AAGTGGTGGG CATCGTGGGA 3420
AGGACGGGCT CTGGGAAGTC CTCCTTGGGC ATGGCTCTCT TCCGCCTGGT GGAGCCCATG 3480
GCAGGCCGGA TTCTCATTGA CGGCGTGGAC ATTTGCAGCA TCGGCCTGGA GGACTTGCGG 3540
TCCAAGCTCT CAGTGATCCC TCAAGATCCA GTGCTGCTCT CAGGAACCAT CAGATTCAAC 3600
CTAGATCCCT TTGACCGTCA CACTGACCAG CAGATCTGGG ATGCCTTGGA GAGGACATTC 3660
CTGACCAAGG CCATCTCAAA GTTCCCCAAA AAGCTGCATA CAGATGTGGT GGAAAACGGT 3720
GGAAACTTCT CTGTGGGGGA GAGGCAGCTG CTCTGCATTG CCAGGGCTGT GCTTCGCAAC 3780
TCCAAGATCA TCCTTATCGA TGAAGCCACA GCCTCCATTG ACATGGAGAC AGACACCCTG 3840
ATCCAGCGCA CAATCCGTGA AGCCTTCCAG GGCTGCACCG TGCTCGTCAT TGCCCACCGT 3900
GTCACCACTG TGCTGAACTG TGACCACATC CTGGTTATGG GCAATGGGAA GGTGGTAGAA 3960
TTTGATCGGC CGGAGGTACT GCGGAAGAAG CCTGGGTCAT TGTTCGCAGC CCTCATGGCC 4020
ACAGCCACTT CTTCACTGAG ATAAGGAGAT GTGGAGACTT CATGGAGGCT GGCAGCTGAG 4080
CTCAGAGGTT CACACAGGTG CAGCTTCGAG GCCCACAGTC TGCGACCTTC TTGTTTGGAG 4140
ATGAGAACTT CTCCTGGAAG CAGGGGTAAA TGTAGGGGGG GTGGGGATTG CTGGATGGAA 4200
ACCCTGGAAT AGGCTACTTG ATGGCTCTCA AGACCTTAGA ACCCCAGAAC CATCTAAGAC 4260
ATGGGATTCA GTGATCATGT GGTTCTCCTT TTAACTTACA TGCTGAATAA TTTTATAATA 4320
AGGTAAAAGC TTATAGTTTT CTGATCTGTG TTAGAAGTGY TGCAAATGCT GTACTGACTT 4380
TGTAAAATAT AAAACTAAGG AAAACTCAAA AAAAAAAAAA AAAAAAA
```

<u>SEQ ID NO:193 BFO6 Protein sequence</u>
Protein Accession #:   NP_115972.1

```
1    11   21   31   41   51
|    |    |    |    |    |
MTRKRTYWVP NSSGGLVNRG IDIGDDMVSG LIYKTYTLQD GPWSQQERNP EAPGRAAVPP  60
WGKYDAALRT MIPFRPKPRF PAPQPLDNAG LFSYLTVSWL TPLMIQSLRS RLDENTIPPL  120
SVHDASDKNV QRLHRLWEEE VSRRGIEKAS VLLVMLRFQR TRLIFDALLG ICFCIASVLG  180
PILIPKILE YSEEQLGNVV HGVGLCFALF LSECVKSLSF SSSWIINQRT AIRFRAAVSS  240
FAFEKLIQFK SVIHITSGEA ISFFTGDVNY LFEGVCYGPL VLITCASLVI CSISSYFIIG  300
YTAFIAILCY LLVFPLAVFM TRMAVKAQHH TSEVSDQRIR VTSEVLTCIK LIKMYTWEKP  360
FAKIIEGMES LTFCSKPGDG MAFSMLASLN LLRLSVFFVP IAVKGLTNSK SAVMRFKKFF  420
LQESPVFYVQ TLQDPSKALV FEEATLSWQQ TCPGIVNGAL ELERNGHASE GMTRPRDALG  480
PEEEGNSLGP ELHKINLVVS KGMMLGVCGN TGSGKSSLLS AILEEMHLLE GSVGVQGSLA  540
YVPQQAWIVS GNIRENILMG GAYDKARYLQ VLHCCSLNRD LELLPFGDMT EIGERGLNLS  600
GGQKQRISLA RAVYSDRQIY LLDDPLSAVD AHVGKHIFEE CIKKTLRGKT VVLVTHQLQY  660
LEFCGQIILL ENGKICENGT HSELMQKKGK YAQLIQKMHK EATSDMLQDT AKIAEKPKVE  720
SQALATSLEE SLNGNAVPEH QLTQEEEMEE GSLSWRVYHH YIQAAGGYMV SCIIFFFVVL  780
IVFLTIFSFW WLSYWLEQGS GTNSSRESNG TMADLGNIAD NPQLSFYQLV YGLNALLLIC  840
VGVCSSGIFT KVTRKASTAL HNKLFNKVFR CPMSFFDTIP IGRLLNCFAG DLEQLDQLLP  900
IFSEQFLVLS LMVIAVLLIV SVLSPYILLM GAIIMVICFI YYMMFKKAIG VFKRLENYSR  960
SPLFSHILNS LQGLSSIHVY GKTEDFISQF KRLTDAQNNY LLLFLSSTRW MALRLEIMTN  1020
LVTLAVALFV AFGISSTPYS FKVMAVNIVL QLASSFQATA RIGLETEAQF TAVERILQYM  1080
KMCVSEAPLH MEGTSCPQGW PQHGEIIFQD YHMKYRDNTP TVLHGINLTI RGHEVVGIVG  1140
RTGSGKSSLG MALFRLVEPM AGRILIDGVD ICSIGLEDLR SKLSVIPQDP VLLSGTIRFN  1200
LDPFDRHTDQ QIWDALERTF LTKAISKFPK KLHTDVVENG GNFSVGERQL LCIARAVLRN  1260
SKIILIDEAT ASIDMETDTL IQRTIREAFQ GCTVLVIAHR VTTVLNCDHI LVMGNGKVVE  1320
FDRPEVLRKK PGSLFAALMA TATSSLR
```

<div align="center">381</div>

SEQ ID NO:194 BHB8 DNA sequence
Nucleic Acid Accession #:   AA983251
Coding sequence: 1-1749 (underlined sequences correspond to start and stop codons)

```
         1          11         21         31         41         51
         |          |          |          |          |          |
    ATGCTGTCTG GCTTCTTGAT GAGTCCCAGT ACCCAGCACA GAGCACAGTA CACTCCCGGA    60
    GGAAAGAAAC TTCCGTGGGA GGCTTCCATC GGTGCGCACA CCTCCCGAGG GCGAGGCAGC   120
    GACCGGGAGA GGGAGAGCCG GCCGGAGGCT GCCGGGCTCC TGTGGGACCG CGCTGCAGCC   180
    GGGGAGGCGG AGAAGGGGAA CCGGGGCGAG CCGCCCGCCT GGATCCGCGC CCAGCAGCAG   240
    CCGCGGCCGC CGCCAGCTGG GCAGGCTCCC GGGACTGCGG CTGGGGGCGC GCAGGACCCT   300
    CGCCTGCGTC CTGGACGTTC CCGGGGGAGG GTCCGGTTGC CAGTGAAACC TCCAGAGGCT   360
    TCCGGACGAC AGCCCCGGGG GCCTTCTGAC TGCATCCCGA GATTTCCATC AGCGAGTGCA   420
    ACTCATAAGG CAGTCCCTAA GGGGACCGGG CCACCGGCTG AGGACGGGGA TGGCTTAGGA   480
    GCTCCTGGAC CTAGGGCCCG GCGTCGTCGC CTCCTGGGCG TCGCGGCAGA GGGGAGTGGC   540
    CCGCGCGGAA AGCGCCGCGG GACAGTCAGT GACGAGGCCC GGGGGTCGCC GGGGCCACGA   600
    CTTCTCGGAG ACCGTCCTGC GCTCTCTGGA GACGCGCTGT CCGCGCCCAG GGTGGTGCCA   660
    TGTGGGGCGC TCGCCGCTCG TCCGTCTCCT CATCCTGGAA CGCCGCTTCG CTCCTGCAGC   720
    TGCTGCTGGC TGCGCTGCTG GCGGCGGGGG CGAGGGCCCA GCGGCGAGTA CTGCCACGGC   780
    TGGCTGGACG CGCAGGGCGT CTGGCGCATC GGCTTCCAGT GTCCCGAGCG CTTCGACGGC   840
    GGCGACGCCA CCATCTGCTG CGGCAGCTGC GCGTTGCGCT ACTGCTGCTC CAGCGCCGAG   900
    GCGCGCCTGG ACCAGGGCGG CTGCGACAAT GACCGCCAGC AGGGCGCTGG CGAGCCTGGC   960
    CGGGCGGACA AAGACGGGCC CCGACGGCTC GGCAGGGCTT CATGTCTTAG GGGTACCCAA  1020
    GGAGACGGCG AGGGTGCGCC CCCACCCGTG AGGGCCTGGC AGCGGTGCTC CCCTGAAGGC  1080
    TCCCCGAAAG GAAGGCAGCT CCTCAGGGCT TTCCCGGGGC TGCTGCCCCG TGCCAGACGC  1140
    CGCGGATTCC CATCTTCTCC ACGCGGCGGC CCCTCTCCCC TGCAGCGGCC CGCCTTGCCC  1200
    ATCTACGTGC CGTTCCTCAT TGTTGGCTCC GTGTTTGTCG CCTTTATCAT CTTGGGGTCC  1260
    CTGGTGGCAG CCTGTTGCTG CAGATGTCTC CGGCCTAAGC AGGATCCCCA GCAGAGCCGA  1320
    GCCCCAGGGG GTAACCGCTT GATGGAGACC ATCCCCATGA TCCCCAGTGC CAGCACCTCC  1380
    CGGGGGTCGT CCTCACGCCA GTCCAGCACA GCTGCCAGTT CCAGCTCCAG CGCCAACTCC  1440
    GGGGCCCGGG CGCCCCCAAC AAGGTCACAG ACCAACTGTT GCTTGCCGGA AGGGACCATG  1500
    AACAACGTGT ATGTCAACAT GCCCACGAAT TTCTCTGTGC TGAACTGTCA GCAGGCCACC  1560
    CAGATTGTGC CACATCAAGG GCAGTATCTG CATCCCCCAT ACGTGGGGTA CACGGTGCAG  1620
    CACGACTCTG TGCCCATGAC AGCTGTGCCA CCTTTCATGG ACGGCCTGCA GCCTGGCTAC  1680
    AGGCAGATTC AGTCCCCCTT CCCTCACACC AACAGTGAAC AGAAGATGTA CCCAGCGGTG  1740
    ACTGTATAAC CGAGAGTCAC TGGTGGGTTC CTTTACTGAA GGGAGACGAA GGCAGGGGTG  1800
    GATTCTCGAG GTGGAAGTCC GCACATGTCG GTGGTATTTA TGGCACGATT CCTTTGGATG  1860
    GCTTCATTTG CCCCCAGACT GTATGAAAAC ATCTCCGAAT TAGCATTTCT GGATATGTTT  1920
    CATCCAGGGT ATCATTGATT TATGATGGAA AACCGGCCTC AGCTGGAGAT GACTGTGATG  1980
    TTGCTGATGG GTGTATAACA AATGCTTGAG TCCGAAGTGC CCTTGAGATA TGGTTGACGA  2040
    AAGAATTTTA TAAACTGATA AATTAAGGAT TTTTATTATG TTGTTATTAT TATTTCTTTT  2100
    TTGTTGTTGA CTGCACAGGA TCAAAATGCC TGTTATCTCC CTTTTACTGG GACTTTTTTT  2160
    TTTTTTTTT TTTTTTTTAA TCAGACAGGG TCTTGCTCTG TTGCCCAGGC TGGAGTGCAG  2220
    TGGTGCGATC TCGGCTCACT GCAACTTCAG CCTCCTGGAT TCAGGCAACA CTCCTGCCTC  2280
    AGCCTCCCAC GTGGCTGGGA TTACAGGTGC CTGCCCCCAT GGCTAATTTT TTGTATTTTT  2340
    TGTAGAGATG GGGTTTCACC ATGTTGGCTG GGCTGGTCTC ACTCTCCTGA CCTCAAGCAA  2400
    TCTGCCTGTC TCAGCCTCCC AAAGTGCTGG GATTACAGGC GTGAGCCACC GCCCCCAGCC  2460
    TGAGCCTTTT TTTTTTTCTA ATGCATCCAA GGTTAAGGGG AAGACGCAAA TAACAGGACT  2520
    ATTCTAAAAG GAAACCTGTT TGAACTCTGT GAGATCAGTC ATCAGTCTCA GTATTCCACA  2580
    GGCACACCTT AATTTCATTG TAAAAAGATA TATATATTTT GTCTATTTTT GTGCTTTTGG  2640
    GGGCCTATTT TGTGCTTTTT TACCTTATGT AGAGATCTTA TTACAAAGTG ATTTTCTACA  2700
    TTAAAAAGAG ACTGAAATAA ATTGTATAGT TACTTAACTA ATGAAGACAT TTCAGAACTC  2760
    TGGGATGATT TTAATCTTGA AGTAGTAGGT GGTATAGTCA TAAAACCATT CATCCCCTTC  2820
    TTGATTGTAT CTTAATTTTC TGGCTTTAAG GTGACATCTG AGAGGTAATG CATTCTTTTT  2880
    TATATTGAAA TCATAAACTA TCACCCGCTG CTTCTCTGAG TTACTTTTAA TTTTGCCTTG  2940
    TGGTTATGGT TTGGCGTTTC CTTCTGTTTG GTTTTCAGAG CCCCATGTCT ATATAGTCCT  3000
    GAGTGCAAGT AATTACTATA CTTGTAAATG AAGATCAGTA TTTCTGCCTA GATCTGATAA  3060
    AAAAATTTTC TTGTCTTAGT TATAAAAATT CAAAGAAATG TGTTACAAAG ATACTTAGTA  3120
    TAGCTCCTCA GCCATAACCT GAGACTTGGG ATGAAATTTA AACCAGATAC GATTTACTTT  3180
    GCAGATCATA AGGCTTTTTA TACTCTTGTT ATCAAAATGG CTTATTTTTC AGGCACTAAG  3240
    GATTGTTAAG AGAAAAGCTT TTCAACGAAG GATTGCCTTT CTTCTCCCAC ACTGTTCTTG  3300
    ATTTCCTCTC TCTTTCAGGC CTCAACAGGC ACTGTATTCA TTGCCAATGT TCCAAATTAT  3360
    CAAATTCATA TGAATTTATT TGTGTGTTCT TTACTTATAT AAAAAAAGAT AACTTTAAGG  3420
    ATGTGCAAGT ACATTTCCAA CTGCTAGCAC AACCAGTATT TTGTAATTAA ACAAATCGCT  3480
    GTATGGTATG GTCTTCTACA CATTTATGTC TATAGATATC TATCGATCAT CTTTCTATTC  3540
    TGTTTCATGA CTGAATAATG TAAAACCAGT GTTGGCAATT GGTATCATCA ATGATACTCA  3600
    TTTTTTAATA ACCAAAGGCA GGGGAAAATC ATTTTACTTA TTAATAAATA TTTTATGATG  3660
    TGAAAAAAAA AAAAAAAAA AAAAAAAAAA                                     
```

SEQ ID NO:195 BHB8 Protein sequence

Protein Accession #: none found

```
     1         11         21         31         41         51
     |          |          |          |          |          |
MLSGFLMSPS TQHRAQYTPG GKKLPWEASI GAHTSRGRGS DRERESRPEA AGLLWDRAAA    60
GEAEKGNRGE PPAWIRAQQQ PRPPPAGQAP GTAAGGAQDP RLRPGRSRGR VRLPVKPPEA   120
SGRQPRGPSD CIPRFPSASA THKAVPKGTG PPAEDGDGLG APGPRARRRR LLGVAAEGSG   180
PRGKRRGTVS DEARGSPGPR LLGDRPALSG DALSAPRVVP CGALAARPSP HPGTPLRSCS   240
CCWLRCWRRG RGPSGEYCHG WLDAQGVWRI GFQCPERFDG GDATICCGSC ALRYCCSSAE   300
ARLDQGGCDN DRQQGAGEPG RADKDGPRRL GRASCLRGTQ GDGEGAPPPV RAWQRCSPEG   360
SPKGRQLLRA FPGLLPRARR RGFPSSPRGG PSPLQRPALP IYVPFLIVGS VFVAFIILGS   420
LVAACCCRCL RPKQDPQQSR APGGNRLMET IPMIPSASTS RGSSSRQSST AASSSSSANS   480
GARAPPTRSQ TNCCLPEGTM NNVYVNMPTN FSVLNCQQAT QIVPHQGQYL HPPYVGYTVQ   540
HDSVPMTAVP PFMDGLQPGY RQIQSPFPHT NSEQKMYPAV TV
```

                                    SEQ ID NO:196 CQA5 DNA SEQUENCE
Nucleic Acid Accession #:   AA088458
                    Coding sequence: 862-1995 (underlined sequences correspond to start and stop codons)

```
     1         11         21         31         41         51
     |          |          |          |          |          |
GCCCTTGGAC ACTGACATGG ACTGAAGGAG TAGAATGGAG CACGAGGACA CTGACATGGA    60
CTGAAGAAAA AGGAGCTGGA GCAGGAGAAG GAGGTGCTGC TGCAGGGTTT GGAGATGATG   120
GCGCGGGGCC GCGACTGGTA CCAGCAGCAG CTGCAACGAG TGCAGGAGCG CCAGCGCCGC   180
CTGGGCCAGA GCAGAGCCAG CGCCGACTTT GGGGCTGCAG GGAGCCCCCG CCCACTGGGG   240
CGGCTACTGC CCAAGGTACA AGAGGTGGCC CGGTGCCTGG GGGAGCTGCT GGCTGCAGCC   300
TGTGCCAGCC GGGCCCTGCC CCCGTCCTCC TCCGGGCCCC CCTGCCCTGC CCTGACGTCC   360
ACCTCACCCC CGGTCTGGCA GCAGCAGACC ATCCTCATGC TGAAGGAGCA GAACCGACTC   420
CTCACCCAGG AGGTGACCGA GAAGAGTGAG CGCATCACGC AGCTGGAGCA GGAGAAGTCG   480
GCGCTCATTA AGCAGCTGTT TGAGGCCCGC GCCCTGAGCC AGCAGGACGG GGGACCTCTG   540
GATTCCACCT TCATCTAGTC CTTGTGGGCC GCGTGGGCCC CCAGGGCCAG CCTGGCACTC   600
AGCCCTTCGA GGGTGGGCGC CCCATCGCAC CCACCCTCTC TGGCTGGAGA CCCCCGGCAG   660
GCCCAGGCAC AGTCCCGGAG TGGGCGCCTT CCTGCCGCCC TTGCCAGATG GGCTCCCCAG   720
GCCTGCCCCC GGCTGGTCCC CGCACCGAGC GCTTGACTCC GTTTKGGCTC CTGGTTGYTG   780
ACATGGGCTG GGGGCTCTCT TGAGTCCGCA TAGTCCGCAG CTACTACTGG CCGCTGTCAG   840
TGGACAGTGG GGTACCCCTC CATGAGTTAG CGTCCCCCCG TTTCCAGCGG TGCCGCCCTG   900
GGTCCATCT TCAGGGAAAG GCACTGCCCA CGCCAGGCTG CACTTCCAAC AACGGGCAGC   960
AGAGGGCGCG GGGCGGCTCC GACGCGGGTC CAAGGGCAGC TTCCCGCTCA ACCAGGGCAC  1020
CAGGACGAGG TGGCTGTAGC TCGGACGGAC GGAAGTAGAT GGAGGGGGTG GGGACGGCCT  1080
GTAAGCGGGG GGTGCCTGCC TGGCTGGGGA GCCCCAGGGA TAGCGGTCGG ACTTCAGGTT  1140
CTGGCCAAGG CTGAGGGACC CTGGCTGCAG CGGATCGGCA CGCCGGGTGG GCGAGAGCTT  1200
GGCCTGCATG TGCCTCCCAC AGACCCTGGG GTGATGGCCT TCCCCCTCTT GGCCGGGACG  1260
TTGCCCCACG TTGAGTCCCA CACAACATCC TGTGAGCCTG GCTCCCCAGG AGGGCCCCCA  1320
GACAGCTCCC AGGCACGTCA TAGGCAAAGC CTGTTTCCCC CGACTCAGGA TTTCCAAGGC  1380
CTGGGGTCCT GCTCACCCCC CTTTGCTCTC ACGCCCAGCC TGTCCCCAGG TTTCAGCTGG  1440
GAGAGGCCAC CTCCCTCAGC CAAGGAAAAC GAGAACCCCC AGGGTACAGG AGGAGGCTGG  1500
GGCAGGTCCC CTTGGGTGTC ACTCCCTCAG CCCCTGCCCA GGCCCACTCC CGCTGGTGCT  1560
GGAGTACGCA CTGGTGGGGG GGCCCTGCTC AGCCCAACCT GGAGGGTCCC AGTGTCACCA  1620
GAACCAGGGG CACGGCAACA GCATCGATGG GTTCTGCAGC CCAGGGCCCC CGATGCGGGG  1680
TCAGTGTGTG TGGGGCGCAG GGCCTCCGAT GCGGGGTCAG TGCGTGGGGG GCGCAGGGCC  1740
CCCGATGCGG GGTCAGTGCG TGGGGGGCGC AGGGCCCCCT CGTGTCCAGG GCACTTTGGT  1800
ACACTGTCCC ACAAGGCACC TGTCTCAGAG GAGGGGCCCT GGCAGGCAGC GTGGCAACTC  1860
CCTTCCGGAG CCCAGCTCCA TGCTAACCTG CCCACAGCAA CCCCACAGAG CCACATTCCC  1920
TGCTGCACCT GGTCTGCAGG GGTGTCCCAG GACAGGCCCA AGTCAGCCCA GCATGCAGCT  1980
GCCCTCCTAC CCTGAAGATG GGAGTGGGCT TTCCAGGGGA CATAAGGATG TCAGGCCTGG  2040
ACCTCCTGGG CAGGAAAGGG TGCAGGTCCT GAGGGCCTGT GCCCCACAGC CCCAGCACCC  2100
AGGTGGACTG CAGCGCAGTG GGTGGGCCAG TGGCAGCCAG GGAGAAGCCC CCCGTCAGCA  2160
GGCTGGGGTC TGCCCACCAG GGCCTCCCCA CGTCTGCCTT TGAGGGTGCC TGCCATGCCC  2220
TGGGGGATCC TGGCATCTTT ACTGGACTGG AAGCAGGAGA CAGAACAGTG TCTGTCCCGG  2280
GGTGACTTCA TCAGGAGACC CCCCACATAG AGCTGGACCC CGCAGCTGAA GCGGAAATGT  2340
GAGACAGGCT GGCACCTCCG GAAAAACTGC CTTTCAGCCT TGGTGTTCCG TGCAAGGTGA  2400
AAAGAAATAG GTCCTCCCAG TTTACAGCTT GAAATCAGGC TAGTGAGTGG CCCTGGAGAC  2460
CACGGGGGA GAATTTAAAG GCCCCGGCTG GCAGGGTCTA GGTGGCTGGC AGAGGCACAT  2520
GCAGACCCTG CCTGGAGCCT GCCCTAGGAC GCTGGGCGGG TCAGTCTCCG TGCAGGATGT  2580
GAGCAGCGTC CCTGGGCTCT ATCCGCGAGG TGCCAGTAGC GTGTGCAGGT ACATACACGT  2640
GCGTGCACAC TGTGATGACA CCCGGAAATG TCTCAGGATG TTGAAATGTG TCCTTGGGGG  2700
CAGAAGTGTC CCCAGTTGAG AATCTGCCCC AGAGGAACAC ACCCACACCA GGCCTCAGGA  2760
TTTTGTGTTG ATCAAGTTCC AAGGAAAAGG AACATCTCAG CCGGGCGTGG TGGTTCACGC  2820
CTGGAATCCC AGCACTTGAG GCCAGGAGTT CCAGAGCAGC CTGGGCAACA CAGTGAGAGA  2880
CCCCATCTCT ACAARAAAAA AAAAAGAAAG AAAGAAAATG AAAGAAAAT CCAGCA  2940
TCATAAACAC CACAAGGAAA CAATACACTA TGAGACCCAG CAGAAGCAAC AGATTGACTC  3000
TAGACCCAGA TACTGAATT ATCAGAGAGA ATATAAAGTA ACAGTGTTTT ATATATCTAA  3060
AGAAATAAAA GAGATTTCTG GAAACATGAA AAAAAA
```

383

SEQ ID NO:197 LBG2 DNA SEQUENCE

Nucleic Acid Accession #:   X63629
Coding sequence: 54-2543 (start and stop codons are underlined)

```
1    11    21    31    41    51
|    |     |     |     |     |
GCGGAACACC GGCCCGCCGT CGCGGCAGCT GCTTCACCCC TCTCTCTGCA GCCATGGGGC   60
TCCCTCGTGG ACCTCTCGCG TCTCTCCTCC TTCTCCAGGT TTGCTGGCTG CAGTGCGCGG   120
CCTCCGAGCC GTGCCGGGCG GTCTTCAGGG AGGCTGAAGT GACCCTTGGAG GCGGGAGGCG   180
CGGAGCAGGA GCCCGGCCAG GCGCTGGGGA AAGTATTCAT GGGCTGCCCT GGGCAAGAGC   240
CAGCTCTGTT TAGCACTGAT AATGATGACT TCACTGTGCG GAATGGCGAG ACAGTCCAGG   300
AAAGAAGGTC ACTGAAGGAA AGGAATCCAT TGAAGATCTT CCCATCCAAA CGTATCTTAC   360
GAAGACACAA GAGAGATTGG GTGGTTGCTC CAATATCTGT CCCTGAAAAT GGCAAGGGTC   420
CCTTCCCCCA GAGACTGAAT CAGCTCAAGT CTAATAAAGA TAGAGACACC AAGATTTTCT   480
ACAGCATCAC GGGGCCGGGG GCAGACAGCC CCCCTGAGGG TGTCTTCGCT GTAGAGAAGG   540
AGACAGGCTG GTTGTTGTTG AATAAGCCAC TGGACCGGGA GGAGATTGCC AAGTATGAGC   600
TCTTTGGCCA CGCTGTGTCA GAGAATGGTG CCTCAGTGGA GGACCCCATG AACATCTCCA   660
TCATCGTGAC CGACCAGAAT GACCACAAGC CCAAGTTTAC CCAGGACACC TTCCGAGGGA   720
GTGTCTTAGA GGGAGTCCTA CCAGGTACTT CTGTGATGCA GGTGACAGCC ACAGATGAGG   780
ATGATGCCAT CTACACCTAC AATGGGGTGG TTGCTTACTC CATCCATAGC CAAGAACCAA   840
AGGACCCACA CGACCTCATG TTCACAATTC ACCGGAGCAC AGGCACCATC AGCGTCATCT   900
CCAGTGGCCT GGACCGGGAA AAAGTCCCTG AGTACACACT GACCATCCAG GCCACAGACA   960
TGGATGGGGA CGGCTCCACC ACCACGGCAG TGGCAGTAGT GGAGATCCTT GATGCCAATG   1020
ACAATGCTCC CATGTTTGAC CCCCAGAAGT ACGAGGCCCA TGTGCCTGAG AATGCAGTGG   1080
GCCATGAGGT GCAGAGGCTG ACGGTCACTG ATCTGGACGC CCCCAACTCA CCAGCGTGGC   1140
GTGCCACCTA CCTTATCATG GGCGGTGACG ACGGGGACCA TTTTACCATC ACCACCCACC   1200
CTGAGAGCAA CCAGGGCATC CTGACAACCA GGAAGGGGTT GGATTTTGAG GCCAAAAACC   1260
AGCACACCCT GTACGTTGAA GTGACCAACG AGGCCCCTTT TGTGCTGAAG CTCCCAACCT   1320
CCACAGCCAC CATAGTGGTC CACGTGGAGG ATGTGAATGA GGCACCTGTG TTTGTCCCAC   1380
CCTCCAAAGT CGTTGAGGTC CAGGAGGGCA TCCCCACTGG GGAGCCTGTG TGTGTCTACA   1440
CTGCAGAAGA CCCTGACAAG GAGAATCAAA AGATCAGCTA CCGCATCCTG AGAGACCCAG   1500
CAGGGTGGCT AGCCATGGAC CCAGACAGTG GGCAGGTCAC AGCTGTGGGC ACCCTCGACC   1560
GTGAGGATGA GCAGTTTGTG AGGAACAACA TCTATGAAGT CATGGTCTTG GCCATGGACA   1620
ATGGAAGCCC TCCCACCACT GGCACGGGAA CCCTTCTGCT AACACTGATT GATGTCAACG   1680
ACCATGGCCC AGTCCCTGAG CCCCGTCAGA TCACCATCTG CAACCAAAGC CCTGTGCGAC   1740
ACGTGCTGAA CATCACGGAC AAGGACCTGT CTCCCCACAC CTCCCCTTTC CAGGCCCAGC   1800
TCACAGATGA CTCAGACATC TACTGGACGG CAGAGGTCAA CGAGGAAGGT GACACAGTGG   1860
TCTTGTCCCT GAAGAAGTTC CTGAAGCAGG ATACATATGA CGTGCACCTT TCTCTGTCTG   1920
ACCATGGCAA CAAAGAGCAG CTGACAGGTGA TCAGGGCCAC TGTGTGCGAC TGCCATGGCC   1980
ATGTCGAAAC CTGCCCTGGA CCCTGGAAAG GAGGTTTCAT CCTCCCTGTG CTGGGGGCTG   2040
TCCTGGCTCT GCTGTTCCTC CTGCTGGTGC TGCTTTTTGTT GGTGAGAAAG AAGCGGAAGA   2100
TCAAGGAGCC CCTCCTACTC CCAGAAGATG ACACCCGTGA CAACGTCTTC TACTATGGCG   2160
AAGAGGGGGG TGGCGAAGAG GACCAGGACT ATGACATCAC CCAGCTCCAC CGAGGTCTGG   2220
AGGCCAGGCC GGAGGTGGTT CTCCGCAATG ACGTGGCACC AACCATCATC CCGACACCCA   2280
TGTACCGTCC TAGGCCAGCC AACCCAGATG AAATCGGCAA CTTTATAATT GAGAACCTGA   2340
AGGCGGCTAA CACAGACCCC ACAGCCCCGC CCTACGACAC CCTCTTGGTG TTCGACTATG   2400
AGGGCAGCGG CTCCGACGCC GCGTCCCTGA GCTCCCTCAC CTCCTCCGCC TCCGACCAAG   2460
ACCAAGATTA CGATTATCTG AACGAGTGGG GCAGCCGCTT CAAGAAGCTG GCAGACATGT   2520
ACGGTGGCGG GGAGGACGAC TAGGCGGCCT GCCTGCAGGG CTGGGGACCA AACGTCAGGC   2580
CACAGAGCAT CTCCAAGGGG TCTCAGTTCC CCCTTCAGCT GAGGACTTCG GAGCTTGTCA   2640
GGAAGTGGCC GTAGCAACTT GGCGGAGACA GGCTATGAGT CTGACGTTAG AGTGGTTGCT   2700
TCCTTAGCCT TTCAGGATGG AGGAATGTGG GCAGTTTGAC TTCAGCACTG AAAACCTCTC   2760
CACCTGGGCC AGGGTTGCCT CAGAGGCCAA GTTTCCAGAA GCCTCTTACC TGCCGTAAAA   2820
TGCTCAACCC TGTGTCCTGG GCCTGGGCCT GCTGTGACTG ACCTACAGTG GACTTTCTCT   2880
CTGGAATGGA ACCTTCTTAG GCCTCCTGGT GCAACTTAAT TTTTTTTTTT AATGCTATCT   2940
TCAAAACGTT AGAGAAAGTT CTTCAAAAGT GCAGCCCAGA GCTGCTGGGC CCACTGGCCG   3000
TCCTGCATTT CTGGTTTCCA GACCCCAATG CCTCCCATTC GGATGGATCT CTGCGTTTTT   3060
ATACTGAGTG TGCCTAGGTT GCCCCCTTATT TTTTATTTTC CCTGTTGCGT TGCTATAGAT   3120
GAAGGGTGAG GACAATCGTG TATATGTACT AGAACTTTTT TATTAAAGAA A
```

SEQ ID NO:198 LBG2 Protein sequence:
Protein Accession #:   CAA45177

```
1    11    21    31    41    51
|    |     |     |     |     |
MGLPRGPLAS LLLLQVCWLQ CAASEPCRAV FREAEVTLEA GGAEQEPGQA LGKVFMGCPG   60
QEPALFSTDN DDFTVRNGET VQERRSLKER NPLKIFPSKR ILRRHKRDWV VAPISVPENG   120
KGPFPQRLNQ LKSNKDRDTK IFYSITGPGA DSPPEGVFAV EKETGWLLLN KPLDREEIAK   180
YELFGHAVSE NGASVEDPMN ISIIVTDQND HKPKFTQDTF RGSVLEGVLP GTSVMQVTAT   240
DEDDAIYTYN GVVAYSIHSQ EPKDPHDLMF TIHRSTGTIS VISSGLDREK VPEYTLTIQA   300
TDMDGDGSTT TAVAVVEILD ANDNAPMFDP QKYEAHVPEN AVGHEVQRLT VTDLDAPNSP   360
AWRATYLIMG GDDGDHFTIT THPESNQGIL TTRKGLDFEA KNQHTLYVEV TNEAPFVLKL   420
PTSTATIVVH VEDVNEAPVF VPPSKVVEVQ EGIPTGEPVC VYTAEDPDKE NQKISYRILR   480
DPAGWLAMDP DSGQVTAVGT LDREDEQFVR NNIYEVMVLA MDNGSPPTTG TGTLLLTLID   540
VNDHGPVPEP RQITICNQSP VRHVLNITDK DLSPHTSPFQ AQLTDDSDIY WTAEVNEEGD   600
TVVLSLKKFL KQDTYDVHLS LSDHGNKEQL TVIRATVCDC HGHVETCPGP WKGGFILPVL   660
GAVLALLFLL LVLLLLVRKK RKIKEPLLLP EDDTRDNVFY YGEEGGGEED QDYDITQLHR   720
```

GLEARPEVVL RNDVAPTIIP TPMYRPRPAN PDEIGNFIIE NLKAANTDPT APPYDTLLVF 780
DYEGSGSDAA SLSSLTSSAS DQDQDYDYLN EWGSRFKKLA DMYGGGEDD

SEQ ID NO:199 OBI5 DNA SEQUENCE

Nucleic Acid Accession #:   NM_012152
Coding sequence: 43-1104 (underlined sequences correspond to start and stop codons)

```
 1          11         21         31         41         51
 |          |          |          |          |          |
CTTCTTTAAA TTTCTTTCTA GGATGTTCAC TTCTTCTCCA CAATGAATGA GTGTCACTAT   60
GACAAGCACA TGGACTTTTT TTATAATAGG AGCAACACTG ATACTGTCGA TGACTGGACA   120
GGAACAAAGC TTGTGATTGT TTTGTGTGTT GGGACGTTTT TCTGCCTGTT TATTTTTTTT   180
TCTAATTCTC TGGTCATCGC GGCAGTGATC AAAAACAGAA AATTTCATTT CCCCTTCTAC   240
TACCTGTTGG CTAATTTAGC TGCTGCCGAT TTCTTCGCTG GAATTGCCTA TGTATTCCTG   300
ATGTTTAACA CAGGCCCAGT TTCAAAAACT TTGACTGTCA ACCGCTGGTT TCTCCGTCAG   360
GGGCTTCTGG ACAGTAGCTT GACTGCTTCC CTCACCAACT TGCTGGTTAT CGCCGTGGAG   420
AGGCACATGT CAATCATGAG GATGCGGGTC CATAGCAACC TGACCAAAAA GAGGGTGACA   480
CTGCTCATTT TGCTTGTCTG GGCCATCGCC ATTTTTATGG GGGCGGTCCC CACACTGGGC   540
TGGAATTGCC TCTGCAACAT CTCTGCCTGC TCTTCCCTGG CCCCCATTTA CAGCAGGAGT   600
TACCTTGTTT TCTGGACAGT GTCCAACCTC ATGGCCTTCC TCATCATGGT TGTGGTGTAC   660
CTGCGGATCT ACGTGTACGT CAAGAGGAAA ACCAACGTCT TGTCTCCGCA TACAAGTGGG   720
TCCATCAGCC GCCGGAGGAC ACCCATGAAG CTAATGAAGA CGGTGATGAC TGTCTTAGGG   780
GCGTTTGTGG TATGCTGGAC CCCGGGCCTG GTGGTTCTGC TCCTCGACGG CCTGAACTGC   840
AGGCAGTGTG GCGTGCAGCA TGTGAAAAGG TGGTTCCTGC TGCTGGCGCT GCTCAACTCC   900
GTCGTGAACC CCATCATCTA CTCCTACAAG GACGAGGACA TGTATGGCAC CATGAAGAAG   960
ATGATCTGCT GCTTCTCTCA GGAGAACCCA GAGAGGCGTC CCTCTCGCAT CCCCTCCACA   1020
GTCCTCAGCA GGAGTGACAC AGGCAGCCAG TACATAGAGG ATAGTATTAG CCAAGGTGCA   1080
GTCTGCAATA AAAGCACTTC CTAAACTCTG GATGCCTCTC GGCCCACCCA GGTGATGACT   1140
GTCTTAGG
```

SEQ ID NO:200 OBI5 Protein sequence:
Protein Accession #:   NP_036284

```
 1          11         21         31         41         51
 |          |          |          |          |          |
MNECHYDKHM DFFYNRSNTD TVDDWTGTKL VIVLCVGTFF CLFIFFSNSL VIAAVIKNRK   60
FHFPFYYLLA NLAAADFFAG IAYVFLMFNT GPVSKTLTVN RWFLRQGLLD SSLTASLTNL   120
LVIAVERHMS IMRMRVHSNL TKKRVTLLIL LVWAIAIFMG AVPTLGWNCL CNISACSSLA   180
PIYSRSYLVF WTVSNLMAFL IMVVVYLRIY VYVKRKTNVL SPHTSGSISR RRTPMKLMKT   240
VMTVLGAFVV CWTPGLVVLL LDGLNCRQCG VQHVKRWFLL LALLNSVVNP IIYSYKDEDM   300
YGTMKKMICC FSQENPERRP SRIPSTVLSR SDTGSQYIED SISQGAVCNK STS
```

SEQ ID NO:201 PAA6 DNA SEQUENCE

Nucleic Acid Accession #:   AA569531
Coding sequence: 1-504 (underlined sequences correspond to start and stop codons)

```
 1          11         21         31         41         51
 |          |          |          |          |          |
ATGACCTACA GTTACTCATT TTTTCAGGCCT GAGTTGATCG TTAATCATCT TAATTATGTT   60
CATTCTGAAG CCAACAGGAG AACCAAGACC AAAACTTTAT TGTCTCTGCT TTCATTTCTT   120
GATGAAACCT CTGGACTAAG CACACATCTT CCTTGTTTAT CTCTCTCAAA GGAGTGTGGA   180
GTGCTTCATC TGGACATCCA CGGGAAGAAG GAAGACATGA GAATCACCCA ACAGTCTTCC   240
CAGCTATACC TGTGGGACAT GGGTGGTTTT ACAATATTTA AGAACCTGTG GATGAGCCTC   300
ATACCCAGAG GGAACAAACG CTCCCCAAAA AGAGTTACGA AAACCATCCT GAGAGATTTT   360
AAGCAGAAGC AAAGTTCAAA GATCCAAGAG GAGAGACGAA GAGAGTCTGC AGGACCAAAC   420
CTCTCTTCAT TCTGGTTTGT GGGGAATGCT GGAAGAGGAG ACAGGCCCCA GATTTGGGCA   480
GGAAGTAAAC AGTTTTCAGG CTGAGGCCAA TCTGAGCAGG AACATTCCAA TATTTCTTCA   540
GCTACGTTGT CCCAGCACTT CACTGGTTAA CCTTTTATGT CCACCATTTG TGGATTTCAC   600
AGCTACTTGT CAATGGTGAA TATTGATCAT CATCATTATC TACTGAGCTG CTACCATATC   660
CCAGCTACTC CTTGCATGTT GTTCATTATT TTCTCAACAC TCAGCATATT TGCAATATGT   720
TATGTAATAT CACAGACAAG GAAACTGAAC GCAGAAATGT TTTATTTCTT GCCAAACATC   780
ACATGAGGAT GAACAATGAA ACCGATTTGA AACCAGGATT GTCTGATTCC AACATCTCTG   840
GGTCCTTTTT CACTCTGATA TGCTGCAATT AAAAAGCCAT TTCTAAGACT GT
```

SEQ ID NO:202 PAA6 Protein sequence:
Protein Accession #:   none found

```
 1          11         21         31         41         51
 |          |          |          |          |          |
MTYSYSFFRP ELIVNHLNYV HSEANRRTKT KTLLSLLSFL DETSGLSTHL PCLSLSKECG   60
VLHLDIHGKK EDMRITQQSS QLYLWDMGGF TIFKNLWMSL IPRGNKRSPK RVTETILRDF   120
KQKQSSKIQE ERRRESAGPN LSSFWFVGNA GRGDRPQIWA GSKQFSG
```

385

SEQ ID NO:203 PAB2 DNA SEQUENCE

Nucleic Acid Accession #: XM_050197
Coding sequence: 310-1971 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
TCACACGTGC CAAGGGGCTG GCTCAGCGGA ACCAGCCTGC ACGCGCTGGC TCCGGGTGAC   60
AGCCGCGCGC CTCGGCCAGG ATCTGAGTGA TGAGACGTGT CCCCACTGAG GTGCCCCACA  120
GCAGCAGGTG TTGAGCATGG GCTGAGAAGC TGGACCGGCA CCAAAGGGCT GGCAGAAATG  180
GGCGCCTGGC TGATTCCTAG GCAGTTGGCG GCAGCAAGGA GGAGAGGCCG CAGCTTCTGG  240
AGCAGAGCCG AGACGAAGCA GTTCTGGAGT GCCTGAACGG CCCCCTGAGC CCTACCCGCC  300
TGGCCCACTA TGGTCCAGAG GCTGTGGGTG AGCCGCCTGC TGCGGCACCG GAAAGCCCAG  360
CTCTTGCTGG TCAACCTGCT AAACCTTTGGC CTGGAGGTGT GTTTGGCCGC AGGCATCACC  420
TATGTGCCGC CTCTGCTGCT GGAAGTGGGG GTAGAGGAGA AGTTCATGAC CATGGTGCTG  480
GGCATTGGTC CAGTGCTGGG CCTGGTCTGT GTCCCGCTCC TAGGCTCAGC CAGTGACCAC  540
TGGCGTGGAC GCTATGGCCG CCGCCGGCCC TTCATCTGGG CACTGTCCTT GGGGCATCCTG  600
CTGAGCCTCT TTCTCATCCC AAGGGCCGGC TGGCTAGCAG GGCTGCTGTG CCCGGATCCC  660
AGGCCCCTGG AGCTGGCACT GCTCATCCTG GGCGTGGGGC TGCTGGACTT CTGTGGCCAG  720
GTGTGCTTCA CTCCACTGGA GGCCCTGCTC TCTGACCTCT TCCGGGACCC GGACCACTGT  780
CGCCAGGCCT ACTCTGTCTA TGCCTTCATG ATCAGTCTTG GGGGCTGCCT GGGCTACCTC  840
CTGCCTGCCA TTGACTGGGA CACCAGTGCC CTGGCCCCCT ACCTGGGCAC CCAGGAGGAG  900
TGCCTCTTTG GCCTGCTCAC CCTCATCTTC CTCACCTGCG TAGCAGCCAC ACTGCTGGTG  960
GCTGAGGAGG CAGCGCTGGG CCCCACCGAG CCAGCAGAAG GGCTGTCGGC CCCCTCCTTG 1020
TCGCCCCACT GCTGTCCATG CCGGGCCCGC TTGGCTTTCC GGAACCTGGG CGCCCTGCTT 1080
CCCCGGCTGC ACCAGCTGTG CTGCCGCATG CCCCGCACCC TGCCGCCGGCT CTTCGTGGCT 1140
GAGCTGTGCA GCTGGATGGC ACTCATGACC TTCACGCTGT TTTACACGGA TTTCGTGGGC 1200
GAGGGGCTGT ACCAGGGCGT GCCCAGAGCT GAGCCGGGCA CCGAGGCCCG GAGACACTAT 1260
GATGAAGGCG TTCGGATGGG CAGCCTGGGG CTGTTCCTGC AGTGCGCCAT CTCCCTGGTC 1320
TTCTCTCTGG TCATGGACCG GCTGGTGCAG CGATTCGGCA CTCGAGCAGT CTATTTGGCC 1380
AGTGTGGCAG CTTTCCCTGT GGCTGCCGGT GCCACATGCC TGTCCCACAG TGTGGCCGTG 1440
GTGACAGCTT CAGCCGCCCT CACCGGGTTC ACCTTCTCAG CCCTGCAGAT CCTGCCCTAC 1500
ACACTGGCCT CCCTCTACCA CCGGGAGAAG CAGGTGTTCC TGCCCAAATA CCGAGGGGAC 1560
ACTGGAGGTG CTAGCAGTGA GGACAGCCTG ATGACCAGCT TCCTGCCAGG CCCTAAGCCT 1620
GGAGCTCCCT TCCCTAATGG ACACGTGGGT GCTGGAGGCA GTGGCCTGCT CCCCACCTCCA 1680
CCCGCGCTCT GCGGGGCCTC TGCCTGTGAT GTCTCCGTAC GTGTGGTGGT GGGTGAGCCC 1740
ACCGAGGCCA GGGTGGTTCC GGGCCGGGGC ATCTGCCTGG ACCTCGCCAT CCTGGATAGT 1800
GCCTTCCTGC TGTCCCAGGT GGCCCCATCC CTGTTTATGG GCTCCATTGT CCAGCTCAGC 1860
CAGTCTGTCA CTGCCTATAT GGTGTCTGCC GCAGGCCTGG GTCTGGTCGC CATTTACTTT 1920
GCTACACAGG TAGTATTTGA CAAGAGCGAC TTGGCCAAAT ACTCAGCCGTA GAAAACTTCC 1980
AGCACATTGG GGTGGAGGGC CTGCCTCACT GGGTCCCAGC TCCCCGCTCC TGTTAGCCCC 2040
ATGGGGCTGC CGGGCTGGCC GCCAGTTTCT GTTGCTGCCA AAGTAATGTG GCTCTCTGCT 2100
GCCACCCTGT GCTGCTGAGG TGCGTAGCTG CACAGCTGGG GGCTGGGGCG TCCCTCTCCT 2160
CTCTCCCCAG TCTCTAGGGC TGCCTGACTG GAGGCCTTCC AAGGGGGTTT CAGTCTGGAC 2220
TTATACAGGG AGGCCAGAAG GGCTCCATGC ACTGGAATGC GGGGACTCTG CAGGTGGATT 2280
ACCCAGGCTC AGGGTTAACA GCTAGCCTCC TAGTTGAGAC ACACCTAGAG AAGGGTTTTT 2340
GGGAGCTGAA TAAACTCAGT CACCTGGTTT CCCATCTCTA AGCCCCTTAA CCTGCAGCTT 2400
CGTTTAATGT AGCTCTTGCA TGGGAGTTTC TAGGATGAAA CACTCCTCCA TGGGATTTGA 2460
ACATATGAAA GTTATTTGTA GGGGAAGAGT CCTGAGGGGC AACACACAAG AACCAGGTCC 2520
CCTCAGCCCC ACAGGCACTG GTCTTTTTTG CTNGANTCCA CCCCCCCCCT CTTTACCCTT 2580
TT
```

SEQ ID NO:204 PAB2 Protein sequence:
Protein Accession #: XP_050197

```
1          11         21         31         41         51
|          |          |          |          |          |
MVQRLWVSRL LRHRKAQLLL VNLLTFGLEV CLAAGITYVP PLLLEVGVEE KFMTMVLGIG   60
PVLGLVCVPL LGSASDHWRG RYGRRRPFIW ALSLGILLSL FLIPRAGWLA GLLCPDPRPL  120
ELALLILGVG LLDFCGQVCF TPLEALLSDL FRDPDHCRQA YSVYAFMISL GGCLGYLLPA  180
IDWDTSALAP YLGTQEECLF GLLTLIFLTC VAATLLVAEE AALGPTEPAE GLSAPSLSPH  240
CCPCRARLAF RNLGALLPRL HQLCCRMPRT LRRLFVAELC SWMALMTFTL FYTDFVGEGL  300
YQGVPRAEPG TEARRHYDEG VRMGSLGLFL QCAISLVFSL VMDRLVQRFG TRAVYLASVA  360
AFPVAAGATC LSHSVAVVTA SAALTGFTFS ALQILPYTLA SLYHREKQVF LPKYRGDTGG  420
ASSEDSLMTS FLPGPKPGAP FPNGHVGAGG SGLLPPPPAL CGASACDVSV RVVVGEPTEA  480
RVVPGRGICL DLAILDSAFL LSQVAPSLFM GSIVQLSQSV TAYMVSAAGL GLVAIYFATQ  540
VVFDKSDLAK YSA
```

SEQ ID NO:205 PAJ3 DNA SEQUENCE

Nucleic Acid Accession #: AK002126
Coding sequence: 1-1593 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGGTTCGCC GGGGGCTGCT TGCGTGGATT TCCCGGGTGG TGGTTTTGCT GGTGCTCCTC   60
TGCTGTGCTA TCTCTGTCCT GTACATGTTG GCCTGCACCC CAAAAGGTGA CGAGGAGCAG  120
CTGGCACTGC CCAGGGCCAA CAGCCCCACG GGGAAGGAGG GGTACCAGGC CGTCCTTCAG  180
GAGTGGGAGG AGCAGCACCG CAACTACGTG AGCAGCCTGA AGCGGCAGAT CGCACAGCTC  240
```

```
AAGGAGGAGC TGCAGGAGAG GAGTGAGCAG CTCAGGAATG GGCAGTACCA AGCCAGCGAT    300
GCTGCTGGCC TGGGTCTGGA CAGGAGCCCC CCAGAGAAAA CCCAGGCCGA CCTCCTGGCC    360
TTCCTGCACT CGCAGGTGGA CAAGGCAGAG GTGAATGCTG GCGTCAAGCT GGCCACAGAG    420
TATGCAGCAG TGCCTTTCGA TAGCTTTACT CTACAGAAGG TGTACCAGCT GGAGACTGGC    480
CTTACCCGCC ACCCCGAGGA GAAGCCTGTG AGGAAGGACA AGCGGGATGA GTTGGTGGAA    540
GCCATTGAAT CAGCCTTGGA GACCCTGAAC AATCCTGCAG AGAACAGCCC CAATCACCGT    600
CCTTACACGG CCTCTGATTT CATAGAAGGG ATCTACCGAA CAGAAAGGGA CAAAGGGACA    660
TTGTATGAGC TCACCTTCAA AGGGGACCAC AAACACGAAT TCAAACGGCT CATCTTATTT    720
CGACCATTCG GCCCCATCAT GAAAGTGAAA AATGAAAAGC TCAACATGGC CAACACGCTT    780
ATCAATGTTA TCGTGCCTCT AGCAAAAAGG GTGGACAAGT TCCGGCAGTT CATGCAGAAT    840
TTCAGGGAGA TGTGCATTGA GCAGGATGGG AGAGTCCATC TCACTGTTGT TTACTTTGGG    900
AAAGAAGAAA TAAATGAAGT CAAAGGAATA CTTGAAAACA CTTCCAAAGC TGCCAACTTC    960
AGGAACTTTA CCTTCATCCA GCTGAATGGA GAATTTTCTC GGGAAAGGG ACTTGATGTT   1020
GGAGCCCGCT TCTGGAAGGG AAGCAACGTC CTTCTCTTTT TCTGTGATGT GGACATCTAC   1080
TTCACATCTG AATTCCTCAA TACGTGTAGG CTGAATACAC AGCCAGGGAA GAAGGTATTT   1140
TATCCAGTTC TTTTCAGTCA GTACAATCCT GGCATAATAT ACGGCCACCA TGATGCAGTC   1200
CCTCCCTTGG AACAGCAGCT GGTCATAAAG AAGGAAACTG GATTTTGGAG AGACTTTGGA   1260
TTTGGGATGA CGTGTCAGTA TCGGTCAGAC TTCATCAATA TAGGTGGGTT TGATCTGGAC   1320
ATCAAAGGCT GGGGCGGAGA GGATGTGCAC CTTTATCGCA AGTATCTCCA CAGCAACCTC   1380
ATAGTGGTAC GGACGCCTGT GCGAGGACTC TTCCACCTCT GGCATGAGAA GCGCTGCATG   1440
GACGAGCTGA CCCCCGAGCA GTACAAGATG TGCATGCAGT CCAAGGCCAT GAACGAGGCA   1500
TCCCACGGCC AGCTGGGCAT GCTGGTGTTC AGGCACGAGA TAGAGGCTCA CCTTCGCAAA   1560
CAGAAACAGA AGACAAGTAG CAAAAAAACA TGA                               1560
```

SEQ ID NO:206 PAJ3 Protein sequence:
Protein Accession #:   NP_060841

```
1          11         21         31         41         51
|          |          |          |          |          |
MVRRGLLAWI SRVVVLLVLL CCAISVLYML ACTPKGDEEQ LALPRANSPT GKEGYQAVLQ    60
EWEEQHRNYV SSLKRQIAQL KEELQERSEQ LRNGQYQASD AAGLGLDRSP PEKTQADLLA   120
FLHSQVDKAE VNAGVKLATE YAAVPFDSFT LQKVYQLETG LTRHPEEKPV RKDKRDELVE   180
AIESALETLN NPAENSPNHR PYTASDFIEG IYRTERDKGT LYELTFKGDH KHEFKRLILF   240
RPFGPIMKVK NEKLNMANTL INVIVPLAKR VDKFRQFMQN FREMCIEQDG RVHLTVVYFG   300
KEEINEVKGI LENTSKAANF RNFTFIQLNG EFSRGKGLDV GARFWKGSNV LLFFCDVDIY   360
FTSEFLNTCR LNTQPGKKVF YPVLFSQYNP GIIYGHHDAV PPLEQQLVIK KETGFWRDFG   420
FGMTCQYRSD FINIGGFDLD IKGWGGEDVH LYRKYLHSNL IVVRTPVRGL FHLWHEKRCM   480
DELTPEQYKM CMQSKAMNEA SHGQLGMLVF RHEIEAHLRK QKQKTSSKKT
```

SEQ ID NO:207 PAJ5 DNA SEQUENCE

Nucleic Acid Accession #:   AF189723
Coding sequence: 1-2712 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGATTCCTG TATTGACATC AAAAAAAGCA AGTGAATTAC CAGTCAGTGA AGTTGCAAGC    60
ATTCTCCAAG CTGATCTTCA GAATGGTCTA AACAAATGTG AAGTTAGTCA TAGGCGAGCC   120
TTTCATGGCT GGAATGAGTT TGATATTAGT GAAGATGAGC CACTGTGGAA GAAGTATATT   180
TCTCAGTTTA AAAATCCCCT TATTATGCTG CTTCTGGCTT CTGCAGTCAT CAGTGTTTTA   240
ATGCATCAGT TTGATGATGC CGTCAGTATC ACTGTGGCAA TACTTATCGT TGTTACAGTT   300
GCCTTTGTTC AGGAATATCG TTCAGAAAAA TCTCTTGAAG AATTGAGTAA ACTTGTGCCA   360
CCAGAATGCC ATTGTGTGCG TGAAGGAAAA TTGGAGCATA CACTTGCCCG AGACTTGGTT   420
CCAGGTGATA CAGTTTGCCT TTCTGTTGGG GATAGAGTTC CTGCTGACTT ACGCTTGTTT   480
GAGGCTGTGG ATCTTTCCAT TGATGAGTCC AGCTTGACAG GTGAGACAAC GCCTTGTTCT   540
AAGGTGACAG CTCCTCAGCC AGCTGCAACT AATGGAGATC TTGCATCGAG AAGTAACATT   600
GCCTTTATGG GAACACTGGT CAGATGTGGC AAAGCAAAGG GTGTTGTCAT TGGAACAGGA   660
GAAAATTCTG AATTTGGGGA GGTTTTTAAA ATGATGCAAG CAGAAGAGGC ACCAAAAACC   720
CCTCTGCAGA AGAGCATGGA CCTCTTAGGA AAACAACTTT CCTTTTACTC CTTTGGTATA   780
ATAGGAATCA TCATGTTGGT TGGCTGGTTA CTGGGAAAAG ATATCCTGGA AATGTTTACT   840
ATTAGTGTAA GTTTGGCTGT AGCAGCAATT CCTGAAGGTC TCCCCATTGT GGTCACAGTG   900
ACGCTAGCTC TTGGTGTTAT GAGAATGGTG AAGAAAAGGG CCATTGTGAA AAAGCTGCCT   960
ATTGTTGAAA CTCTGGGCTG CTGTAATGTG ATTTGTTCAG ATAAAACTGG AACACTGACG  1020
AAGAATGAAA TGACTGTTAC TCACATATTT ACTTCAGATG GTCTGCATCC TGAGGTTACT  1080
GGAGTTGGCT ATAATCAATT TGGGGAAGTG ATTGTTGATG GTGATGTTGT TCATGGATTC  1140
TATAACCCAG CTGTTAGCAG AATTGTTGAG GCGGGCTGTG TGTGCAATGA TGCTGTAATT  1200
AGAAACAATA CTCTAATGGG GAAGCCAACA GAAGGGGCCT TAATTGCTCT TGCAATGAAG  1260
ATGGGTCTTG ATGGACTTCA ACAAGACTAC ATCAGAAAAG CTGAATACCC TTTTAGCTCT  1320
GAGCAAAAGT GGATGGCTGT TAAGTGTGTA CACCGAACAC AGCAGGACAG ACCAGAGATT  1380
TGTTTTATGA AAGGTGCTTA CGAACAAGTA ATTAAGTACT GTACTACATA CCAGAGCAAA  1440
GGGCAGACCT TGACACTTAC TCAGCAGCAG AGAGATGTGT ACCAACAAGA GAAGGCACGC  1500
ATGGGCTCAG CGGGACTCAG AGTTCTTGCT TTGGCTTCTG GTCCTGAACT GGGACAGCTG  1560
ACATTTCTTG GCTTGGTGGG AATCATTGAT CCACCTAGAA CTGGTGTGAA AGAAGCTGTT  1620
ACAACACTCA TTGCCTCAGG AGTATCAATA AAAATGATTA CTGGAGATTC ACAGGAGACT  1680
GCAGTTGCAA TCGCCAGTCG TCTGGGATTG TATTCCAAAA CTTCCCAGTC AGTCTCAGGA  1740
GAAGAAATAG ATGCAATGGA TGTTCAGCAG CTTTCACAAA TAGTACCAAA GGTTGCAGTA  1800
TTTTACAGAG CTAGCCCAAG GCACAAGATG AAAATTATTA AGTCGCTACA GAAGAACGGT  1860
TCAGTTGTAG CCATGACAGG AGATGGAGTA AATGATGCAG TTGCTCTGAA GGCTGCAGAC  1920
```

387

```
ATTGGAGTTG CGATGGGCCA GACTGGTACA GATGTTTGCA AAGAGGCAGC AGACATGATC  1980
CTAGTGGATG ATGATTTTCA AACCATAATG TCTGCAATCG AAGAGGGTAA AGGGATTTAT  2040
AATAACATTA AAAATTTCGT TAGATTCCAG CTGAGCACGA GTATAGCAGC ATTAACTTTA  2100
ATCTCATTGG CTACATTAAT GAACTTTCCT AATCCTCTCA ATGCCATGCA GATTTTGTGG  2160
ATCAATATTA TTATGGATGG ACCCCCAGCT CAGAGCCTTG GAGTAGAACC AGTGGATAAA  2220
GATGTCATTC GTAAACCTCC TCGCAACTGG AAAGACAGCA TTTTGACTAA AAACTTGATA  2280
CTTAAAATAC TTGTTTCATC AATAATCATT GTTTGTGGGA CTTTGTTTGT CTTCTGGCGT  2340
GAGCTACGAG ACAATGTGAT TACACCTCGA GACACAACGA TGACCTTCAC ATGCTTTGTG  2400
TTTTTTGACA TGTTCAATGC ACTAAGTTCC AGATCCCAGA CCAAGTCTGT GTTTGAGATT  2460
GGACTCTGCA GTAATAGAAT GTTTTGCTAT GCAGTTCTTG GATCCATCAT GGGACAATTA  2520
CTAGTTATTT ACTTTCCTCC GCTTCAGAAG GTTTTTCAGA CTGAGAGCCT AAGCATACTG  2580
GATCTGTTGT TTCTTTTGGG TCTCACCTCA TCAGTGTGCA TAGTGGCAGA AATTATAAAG  2640
AAGGTTGAAA GGAGCAGGGA AAAAGATCCAG AAGCATGTTA GTTCGACATC ATCATCTTTT  2700
CTTGAAGTAT GA
```

<u>SEQ ID NO:208 PAJ5 Protein sequence:</u>
Protein Accession #:   AAF27813

```
1          11         21         31         41         51
|          |          |          |          |          |
MIPVLTSKKA SELPVSEVAS ILQADLQNGL NKCEVSHRRA FHGWNEFDIS EDEPLWKKYI  60
SQFKNPLIML LLASAVISVL MHQFDDAVSI TVAILIVVTV AFVQEYRSEK SLEELSKLVP  120
PECHCVREGK LEHTLARDLV PGDTVCLSVG DRVPADLRLF EAVDLSIDES SLTGETTPCS  180
KVTAPQPAAT NGDLASRSNI AFMGTLVRCG KAKGVVIGTG ENSEFGEVFK MMQAEEAPKT  240
PLQKSMDLLG KQLSFYSFGI IGIIMLVGWL LGKDILEMFT ISVSLAVAAI PEGLPIVVTV  300
TLALGVMRMV KKRAIVKKLP IVETLGCCNV ICSDKTGTLT KNEMTVTHIF TSDGLHAEVT  360
GVGYNQFGEV IVDGDVVHGF YNPAVSRIVE AGCVCNDAVI RNNTLMGKPT EGALIALAMK  420
MGLDGLQQDY IRKAEYPFSS EQKWMAVKCV HRTQQDRPEI CFMKGAYEQV IKYCTTYQSK  480
GQTLTLTQQQ RDVYQQEKAR MGSAGLRVLA LASGPELGQL TFLGLVGIID PPRTGVKEAV  540
TTLIASGVSI KMITGDSQET AVAIASRLGL YSKTSQSVSG EEIDAMDVQQ LSQIVPKVAV  600
FYRASPRHKM KIIKSLQKNG SVVAMTGDGV NDAVALKAAD IGVAMGQTGT DVCKEAADMI  660
LVDDDFQTIM SAIEEGKGIY NNIKNFVRFQ LSTSIAALTL ISLATLMNFP NPLNAMQILW  720
INIIMDGPPA QSLGVEPVDK DVIRKPPRNW KDSILTKNLI LKILVSSIII VCGTLFVFWR  780
ELRDNVITPR DTTMTFTCFV FFDMFNALSS RSQTKSVFEI GLCSNRMFCY AVLGSIMGQL  840
LVIYFPPLQK VFQTESLSIL DLLFLLGLTS SVCIVAEIIK KVERSREKIQ KHVSSTSSSF  900
LEV
```

SEQ ID NO:209 PAV4 VARIANT 1 DNA SEQUENCE
Nucleic Acid Accession #:   N62096
Coding sequence: 1-1284 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGGGCTACC AGAGGCAGGA GCCTGTCATC CCGCCGCAGA GAGGATTGCC TTATTCAATG  60
AAGCAAGCTG GGTTTCCTTT GGGAATATTG CTTTTATTCT GGGTTTCATA TGTTACAGAC  120
TTTTTCCCTTG TTTTATTGAT AAAAGGAGGG GCCCTCTCTG GAACAGATAC CTACCAGTCT  180
TTGGTCAATA AAACTTTCGG CTTTCCAGGG TATCTGCTCC TCTCTGTTCT TCAGTTTTTG  240
TATCCTTTTA TAGCAATGAT AAGTTACAAT ATAATAGCTG GAGATACTTT GAGCAAAGTT  300
TTTCAAAGAA TCCCAGGAGT TGATCCTGAA AACGTGTTTA TTGGTCGCCA CTTCATTATT  360
GGACTTTCCA CAGTTACCTT TACTCTGCCT TTATCCTTGT ACCGAAATAT AGCAAAGCTT  420
GGAAAGGTCT CCCTCATCTC TACAGGTTTA ACAACTCTGA TTCTTGGAAT TGTAATGGCA  480
AGGGCAATTT CACTGGGTCC ACACATACCA AAAACAGAAG ACGCTTGGGT ATTTGCAAAG  540
CCCAATGCCA TTCAAGCGGT CGGGGTTATG TCTTTTGCAT TTATTTGCCA CCATAACTCC  600
TTCTTAGTTT ACAGTTCTCT AGAAGAACCC ACAGTAGCTA AGTGGTCCCG CCTTATCCAT  660
ATGTCCATCG TGATTTCTGT ATTTATCTGT ATATTCTTTG CTACATGTGG ATACTTGACA  720
TTTACTGGCT TCACCCAAGG GGACTTATTT GAAAATTACT GCAGAAATGA TGACCTGGTA  780
ACATTTGGAA GATTTTGTTA TGGTGTCACT GTCATTTTGA CATACCCTAT GGAATGCTTT  840
GTGACAAGAG AGGTAATTGC CAATGTGTTT TTTGGTGGGA ATCTTTCATC GGTTTTCCAC  900
ATTGTTGTAA CAGTGATGGT CATCACTGTA GCCACGCTTG TGTCATTGCT GATTGATTGC  960
CTCGGGATAG TTCTAGAACT CAATGGTGTG CTCTGTGCAA CTCCCCTCAT TTTTATCATT  1020
CCATCAGCCT GTTATCTGAA ACTGTCTGAA GAACCAAGGA CACACTCCGA TAAGATTATG  1080
TCTTGTGTCA TGCTTCCCAT TGGTGCTGTG GTGATGGTTT TTGGATTCGT CATGGCTATT  1140
ACAAATACTC AAGACTGCAC CCATGGGCAG GAAATGTTCT ACTGCTTTCC TGACAATTTC  1200
TCTCTCACAA ATACCTCAGA GTCTCATGTT CAGCAGACAA CACAACTTTC TACTTTAAAT  1260
ATTAGTATCT TTCAACTCGA GTAA
```

<u>SEQ ID NO:210 PAV4 Variant 1 Protein sequence:</u>
Protein Accession #:   none found

```
1          11         21         31         41         51
|          |          |          |          |          |
MGYQRQEPVI PPQRGLPYSM KQAGFPLGIL LLFWVSYVTD FSLVLLIKGG ALSGTDTYQS  60
LVNKTFGFPG YLLLSVLQFL YPFIAMISYN IIAGDTLSKV FQRIPGVDPE NVFIGRHFII  120
GLSTVTFTLP LSLYRNIAKL GKVSLISTGL TTLILGIVMA RAISLGPHIP KTEDAWVFAK  180
PNAIQAVGVM SFAFICHHNS FLVYSSLEEP TVAKWSRLIH MSIVISVFIC IFFATCGYLT  240
FTGFTQGDLF ENYCRNDDLV TFGRFCYGVT VILTYPMECF VTREVIANVF FGGNLSSVFH  300
IVVTVMVITV ATLVSLLIDC LGIVLELNGV LCATPLIFII PSACYLKLSE EPRTHSDKIM  360
SCVMLPIGAV VMVFGFVMAI TNTQDCTHGQ EMFYCFPDNF SLTNTSESHV QQTTQLSTLN  420
```

388

ISIFQLE

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGGGCTACC AGAGGCAGGA GCCTGTCATC CCGCCGCAGT TTTCCCTTGT TTTATTGATA   60
AAAGGAGGGG CCCTCTCTGG AACAGATACC TACCAGTCTT TGGTCAATAA AACTTTCGGC  120
TTTCCAGGGT ATCTGCTCCT CTCTGTTCTT CAGTTTTTGT ATCCTTTTAT AGCAATGATA  180
AGTTACAATA TAATAGCTGG AGATACTTTG AGCAAAGTTT TTCAAAGAAT CCCAGGAGTT  240
GATCCTGAAA ACGTGTTTAT TGGTCGCCAC TTCATTATTG GACTTTCCAC AGTTACCTTT  300
ACTCTGCCTT TATCCTTGTA CCGAAATATA GCAAAGCTTG GAAAGGTCTC CCTCATCTCT  360
ACAGGTTTAA CAACTCTGAT TCTTGGAATT GTAATGGCAA GGGCAATTTC ACTGGGTCCA  420
CACATACCAA AAACAGAAGA CGCTTGGGTA TTTGCAAAGC CCAATGCCAT TCAAGCGGTC  480
GGGGTTATGT CTTTTGCATT TATTTGCCAC CATAACTCCT TCTTAGTTTA CAGTTCTCTA  540
GAAGAACCCA CAGTAGCTAA GTGGTCCCGC CTTATCCATA TGTCCATCGT GATTTCTGTA  600
TTTATCTGTA TATTCTTTGC TACATGTGGA TACTTGACAT TTACTGGCTT CACCCAAGGG  660
GACTTATTTG AAAATTACTG CAGAAATGAT GACCTGGTAA CATTTGGAAG ATTTTGTTAT  720
GGTGTCACTG TCATTTTGAC ATACCCTATG GAATGCTTTG TGACAAGAGA GGTAATTGCC  780
AATGTGTTTT TTGGTGGGAA TCTTTCATCG GTTTTCCACA TTGTTGTAAC AGTGATGGTC  840
ATCACTGTAG CCACGCTTGT GTCATTGCTG ATTGATTGCC TCGGGATAGT TCTAGAACTC  900
AATGGTGTGC TCTGTGCAAC TCCCCTCATT TTTATCATTC CATCAGCCTG TTATCTGAAA  960
CTGTCTGAAG AACCAAGGAC ACACTCCGAT AAGATTATGT CTTGTGTCAT GCTTCCCATT 1020
GGTGCTGTGG TGATGGTTTT TGGATTCGTC ATGGCTATTA CAAATACTCA AGACTGCACC 1080
CATGGGCAGG AAATGTTCTA CTGCTTTCCT GACAATTTCT CTCTCACAAA TACCTCAGAG 1140
TCTCATGTTC AGCAGACAAC ACAACTTTCT ACTTTAAATA TTAGTATCTT TCAACTCGAG 1200
TAA
```

```
1          11         21         31         41         51
|          |          |          |          |          |
MGYQRQEPVI PPQFSLVLLI KGGALSGTDT YQSLVNKTFG FPGYLLLSVL QFLYPFIAMI   60
SYNIIAGDTL SKVFQRIPGV DPENVFIGRH FIIGLSTVTF TLPLSLYRNI AKLGKVSLIS  120
TGLTTLILGI VMARAISLGP HIPKTEDAWV FAKPNAIQAV GVMSFAFICH HNSFLVYSSL  180
EEPTVAKWSR LIHMSIVISV FICIFFATCG YLTFTGFTQG DLFENYCRND DLVTFGRFCY  240
GVTVILTYPM ECFVTREVIA NVFFGGNLSS VFHIVVTVMV ITVATLVSLL IDCLGIVLEL  300
NGVLCATPLI FIIPSACYLK LSEEPRTHSD KIMSCVMLPI GAVVMVFGFV MAITNTQDCT  360
HGQEMFYCFP DNFSLTNTSE SHVQQTTQLS TLNISIFQLE
```

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGGGCTACC AGAGGCAGGA GCCTGTCATC CCGCCGCAGG TCAATAAAAC TTTCGGCTTT   60
CCAGGGTATC TGCTCCTCTC TGTTCTTCAG TTTTTGTATC CTTTTATAGC AATGATAAGT  120
TACAATATAA TAGCTGGAGA TACTTTGAGC AAAGTTTTTC AAAGAATCCC AGGAGTTGAT  180
CCTGAAAACG TGTTTATTGG TCGCCACTTC ATTATTGGAC TTTCCACAGT TACCTTTACT  240
CTGCCTTTAT CCTTGTACCG AAATATAGCA AAGCTTGGAA AGGTCTCCCT CATCTCTACA  300
GGTTTAACAA CTCTGATTCT TGGAATTGTA ATGGCAAGGG CAATTTCACT GGGTCCACAC  360
ATACCAAAAA CAGAAGACGC TTGGGTATTT GCAAAGCCCA ATGCCATTCA AGCGGTCGGG  420
GTTATGTCTT TTGCATTTAT TTGCCACCAT AACTCCTTCT TAGTTTACAG TTCTCTAGAA  480
GAACCCACAG TAGCTAAGTG GTCCCGCCTT ATCCATATGT CCATCGTGAT TTCTGTATTT  540
ATCTGTATAT TCTTTGCTAC ATGTGGATAC TTGACATTTA CTGGCTTCAC CCAAGGGGAC  600
TTATTTGAAA ATTACTGCAG AAATGATGAC CTGGTAACAT TTGGAAGATT TTGTTATGGT  660
GTCACTGTCA TTTTGACATA CCCTATGGAA TGCTTTGTGA CAAGAGAGGT AATTGCCAAT  720
GTGTTTTTTG GTGGGAATCT TTCATCGGTT TTCCACATTG TTGTAACAGT GATGGTCATC  780
ACTGTAGCCA CGCTTGTGTC ATTGCTGATT GATTGCCTCG GGATAGTTCT AGAACTCAAT  840
GGTGTGCTCT GTGCAACTCC CCTCATTTTT ATCATTCCAT CAGCCTGTTA TCTGAAACTG  900
TCTGAAGAAC AAGGACACAC TCCGATAAG ATTATGTCTT GTGTCATGCT TCCCATTGGT  960
GCTGTGGTGA TGGTTTTTGG ATTCGTCATG GCTATTACAA ATACTCAAGA CTGCACCCAT 1020
GGGCAGGAAA TGTTCTACTG CTTTCCTGAC AATTTCTCTC TCACAAATAC CTCAGAGTCT 1080
CATGTTCAGC AGACAACACA ACTTTCTACT TTAAATATTA GTATCTTTCA ACTCGAGTAA 1140
```

```
1          11         21         31         41         51
|          |          |          |          |          |
```

389

```
MGYQRQEPVI PPQVNKTFGF PGYLLLSVLQ FLYPFIAMIS YNIIAGDTLS KVFQRIPGVD    60
PENVFIGRHF IIGLSTVTFT LPLSLYRNIA KLGKVSLIST GLTTLILGIV MARAISLGPH   120
IPKTEDAWVF AKPNAIQAVG VMSFAFICHH NSFLVYSSLE EPTVAKWSRL IHMSIVISVF   180
ICIFFATCGY LTFTGFTQGD LFENYCRNDD LVTFGRFCYG VTVILTYPME CFVTREVIAN   240
VFFGGNLSSV FHIVVTVMVI TVATLVSLLI DCLGIVLELN GVLCATPLIF IIPSACYLKL   300
SEEPRTHSDK IMSCVMLPIG AVVMVFGFVM AITNTQDCTH GQEMFYCFPD NFSLTNTSES   360
HVQQTTQLST LNISIFQLE
```

Nucleic Acid Accession #:   N62096
Coding sequence: 1-1389 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGGGCTACC AGAGGCAGGA GCCTGTCATC CCGCCGCAGA GAGATTTAGA TGACAGAGAA    60
ACCCTTGTTT CTGAACATGA GTATAAAGAG AAAACCTGTC AGTCTGCTGC TCTTTTTAAT   120
GTTGTCAACT CGATTATAGG ATCTGGTATA ATAGGATTGC CTTATTCAAT GAAGCAAGCT   180
GGGTTTCCTT TGGGAATATT GCTTTTATTC TGGGTTTCAT ATGTTACAGA CTTTTCCCTT   240
GTTTTATTGA TAAAAGGAGG GGCCCTCTCT GGAACAGATA CCTACCAGTC TTTGGTCAAT   300
AAAACTTTCG GCTTTCCAGG GTATCTGCTC CTCTCTGTTC TTCAGTTTTT GTATCCTTTT   360
ATAGCAATGA TAAGTTACAA TATAATAGCT GGAGATACTT TGAGCAAAGT TTTTCAAAGA   420
ATCCCAGGAG TTGATCCTGA AAACGTGTTT ATTGGTCGCC ACTTCATTAT TGGACTTTCC   480
ACAGTTACCT TTACTCTGCC TTTATCCTTG TACCGAAATA TAGCAAAGCT TGGAAAGGTC   540
TCCCTCATCT CTACAGGTTT AACAACTCTG ATTCTTGGAA TTGTAATGGC AAGGGCAATT   600
TCACTGGGTC CACACATACC AAAAACAGAA GACGCTTGGG TATTTGCAAA GCCCAATGCC   660
ATTCAAGCGG TCGGGGTTAT GTCTTTTGCA TTTATTTGCC ACCATAACTC CTTCTTAGTT   720
TACAGTTCTC TAGAAGAACC CACAGTAGCT AAGTGGTCCC GCCTTATCCA TATGTCCATC   780
GTGATTTCTG TATTTATCTG TATATTCTTT GCTACATGTG GATACTTGAC ATTTACTGGC   840
TTCACCCAAG GGGACTTATT TGAAAATTAC TGCAGAAATG ATGACCTGGT AACATTTGGA   900
AGATTTTGTT ATGGTGTCAC TGTCATTTTG ACATACCCTA TGGAATGCTT TGTGACAAGA   960
GAGGTAATTG CCAATGTGTT TTTTGGTGGG AATCTTTCAT CGGTTTTCCA CATTGTTGTA  1020
ACAGTGATGG TCATCACTGT AGCCACGCTT GTGTCATTGC TGATTGATTG CCTCGGGATA  1080
GTTCTAGAAC TCAATGGTGT GCTCTGTGCA ACTCCCCTCA TTTTTATCAT TCCATCAGCC  1140
TGTTATCTGA AACTGTCTGA AGAACCAAGG ACACACTCCG ATAAGATTAT GTCTTGTGTC  1200
ATGCTTCCCA TTGGTGCTGT GGTGATGGTT TTTGGATTCG TCATGGCTAT TACAAATACT  1260
CAAGACTGCA CCCATGGGCA GGAAATGTTC TACTGCTTTC CTGACAATTT CTCTCTCACA  1320
AATACCTCAG AGTCTCATGT TCAGCAGACA ACACAACTTT CTACTTTAAA TATTAGTATC  1380
TTTCAATGA
```

SEQ ID NO:216 PAV4 Variant 4 Protein sequence:
Protein Accession #:   none found

```
1          11         21         31         41         51
|          |          |          |          |          |
MGYQRQEPVI PPQRDLDDRE TLVSEHEYKE KTCQSAALFN VVNSIIGSGI IGLPYSMKQA    60
GFPLGILLLF WVSYVTDFSL VLLIKGGALS GTDTYQSLVN KTFGFPGYLL LSVLQFLYPF   120
IAMISYNIIA GDTLSKVFQR IPGVDPENVF IGRHFIIGLS TVTFTLPLSL YRNIAKLGKV   180
SLISTGLTTL ILGIVMARAI SLGPHIPKTE DAWVFAKPNA IQAVGVMSFA FICHHNSFLV   240
YSSLEEPTVA KWSRLIHMSI VISVFICIFF ATCGYLTFTG FTQGDLFENY CRNDDLVTFG   300
RFCYGVTVIL TYPMECFVTR EVIANVFFGG NLSSVFHIVV TVMVITVATL VSLLIDCLGI   360
VLELNGVLCA TPLIFIIPSA CYKLSEEPR THSDKIMSCV MLPIGAVVMV FGFVMAITNT   420
QDCTHGQEMF YCFPDNFSLT NTSESHVQQT TQLSTLNISI FQ
```

Nucleic Acid Accession #:   NM_017636
Coding sequence: 1-3501 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGGAGGATG CCTTCGGGGC AGCCGTGGTG ACCGTGTGGG ACACGCGATGC ACACACCACG    60
GAGAAGCCCA CCGATGCCTA CGGAGAGCTG GACTTCACGG GGGCCGGCCG CAAGCACAGC   120
AATTTCCTCC GGCTCTCTGA CCGAACGGAT CCAGCTGCAG TTTATAGTCT GGTCACACGC   180
ACATGGGGCT TCCGTGCCCC GAACCTGGTG GTGTCAGTGC TGGGGGGATC GGGGGGCCCC   240
GTCCTCCAGA CCTGGCTGCA GGACCTGCTG CGTCGTGGGC TGGTGCGGGC TGCCCAGAGC   300
ACAGGAGCCT GGATTGTCAC TGGGGGTCTG CACACAGGGCA TCGGCCGGCA TGTTGGTGTG   360
GCTGTACGGG ACCATCAGAT GGCCAGCACT GGGGGCACCA AGGTGGTGGC CATGGGTGTG   420
GCCCCCTGGG GTGTGGTCCG GAATAGAGAC ACCCTCATCA ACCCCAAGGG CTCGTTCCCT   480
GCGAGGTACC GGTGGCGCGG TGACCCGGAG GACGGGGTCC AGTTTCCCCT GGACTACAAC   540
TACTCGGCCT TCTTCCTGGT GGACGACGGC ACACACGGCT GCCTGGGGGG CGAGAACCGC   600
TTCCGCTTGC GCCTGGAGTC CTACATCTCA CAGCAGAAGA CGGGCGTGGG AGGGACTGGA   660
ATTGACATCC CTGTCCTGCT CCTCCTGATT GATGGTGATG AGAAGATGTT GACGCGAATA   720
GAGAACGCCA CCCAGGCTCA GCTCCCATGT CTCCTCGTGG CTGGCTCAGG GGGAGCTGCG   780
GACTGCCTGG CGGAGACCCT GGAAGACACT CTGGCCCCAG GGAGTGGGGG AGCCAGGCAA   840
GGCGAAGCCC GAGATCGAAT CAGGCGTTTC TTTCCCAAAG GGGACCTTGA GGTCCTGCAG   900
GCCCAGGTGG AGAGGATTAT GACCCGGAAG GAGCTCCTGA CAGTCTATTC TTCTGAGGAT   960
```

```
GGGTCTGAGG AATTCGAGAC CATAGTTTTG AAGGCCCTTG TGAAGGCCTG TGGGAGCTCG   1020
GAGGCCTCAG CCTACCTGGA TGAGCTGCGT TTGGCTGTGG CTTGGAACCG CGTGGACATT   1080
GCCCAGAGTG AACTCTTTCG GGGGGACATC CAATGGCAGT CCTTCCATCT CGAAGCTTCC   1140
CTCATGGACG CCCTGCTGAA TGACCGGCCT GAGTTCGTGC GCTTGCTCAT TTCCCACGGC   1200
CTCAGCCTGG GCCACTTCCT GACCCCGATG CGCCTGGCCC AACTCTACAG CGCGGCGCCC   1260
TCCAACTCGC TCATCCGCAA CCTTTTGGAC CAGGCGTCCC ACAGCGCAGG CACCAAAGCC   1320
CCAGCCCTAA AAGGGGGAGC TGCGGAGCTC CGGCCCCCTG ACGTGGGGCA TGTGCTGAGG   1380
ATGCTGCTGG GGAAGATGTG CGCGCCGAGG TACCCCTCCG GGGGCGCCTG GGACCCTCAC   1440
CCAGGCCAGG GCTTCGGGGA GAGCATGTAT CTGCTCTCGG ACAAGGCCAC CTCGCCGCTC   1500
TCGCTGGATG CTGGCCTCGG GCAGGCCCCC TGGAGCGACC TGCTTCTTTG GGCACTGTTG   1560
CTGAACAGGG CACAGATGGC CATGTACTTC TGGGAGATGG GTTCCAATGC AGTTTCCTCA   1620
GCTCTTGGGG CCTGTTTGCT GCTCCGGGTG ATGGCACGCC TGGAGCCTGA CGCTGAGGAG   1680
GCAGCACGGA GGAAAGACCT GGCCGTTCAAG TTTGAGGGGA TGGGCGTTGA CCTCTTTGGC   1740
GAGTGCTATC GCAGCAGTGA GGTGAGGGCT GCCCGCCTCC TCCTCCGTCG CTGCCCGCTC   1800
TGGGGGGATG CCACTTGCCT CCAGCTGGCC ATGCAAGCTG ACGCCCGTGC CTTCTTTGCC   1860
CAGGATGGGG TACAGTCTCT GCTGACACAG AAGTGGTGGG GAGATATGGC CAGCACTACA   1920
CCCATCTGGG CCCTGGTTCT CGCCTTCTTT TGCCCTCCAC TCATCTACAC CCGCCTCATC   1980
ACCTTCAGGA AATCAGAAGA GGAGCCCACA CGGGAGGAGC TAGAGTTTGA CATGGATAGT   2040
GTCATTAATG GGGAAGGGCC TGTCGGGACG GCGGACCCAG CCGAGAAGAC GCCGCTGGGG   2100
GTCCCGCGCC AGTCGGGCCG TCCGGGTTGC TGCGGGGGCC GCTGCGGGGG GCGCCGGTGC   2160
CTACGCCGCT GGTTCCACTT CTGGGGCGCG CCGGTGACCA TCTTCATGGG CAACGTGGTC   2220
AGCTACCTGC TGTTCCTGCT GCTTTTCTCG CGGGTGCTGC TCGTGGATTT CCAGCCGGCG   2280
CCGCCCGGCT CCCTGGAGCT GCTGCTCTAT TTCTGGGCTT TCACGCTGCT GTGCGAGGAA   2340
CTGCGCCAGG GCCTGAGCGG AGGCGGGGGC AGCCTCGCCA GCGGGGGGCCC CGGGCCTGGC   2400
CATGCCTCAC TGAGCCAGCG CCTGCGCCTC TACCTCGCCG ACAGCTGGAA CCAGTGCGAC   2460
CTAGTGGCTC TCACCTGCTT CCTCCTGGGC GTGGGCTGCC GGCTGACCCC GGGTTTGTAC   2520
CACCTGGGCC GCACTGTCCT CTGCATCGAC TTCATGGTTT TCACGGTGCG GCTGCTTCAC   2580
ATCTTCACGG TCAACAAACA GCTGGGGCCC AAGATCGTCA TCGTGAGCAA GATGATGAAG   2640
GACGTGTTCT TCTTCCTCTT CTTCCTCGGC GTGTGGCTGG TAGCCTATGG CGTGGCCACG   2700
GAGGGGCTCC TGAGGCCACG GGACAGTGAC TTCCCAAGTA TCCTGCGCCG CGTCTTCTAC   2760
CGTCCCTACC TGCAGATCTT CGGGCAGATT CCCCAGGAGG ACATGGACGT GGCCCTCATG   2820
GAGCACAGCA ACTGCTCGTC GGAGCCCGGC TTCTGGGCAC ACCCTCCTGG GGCCCAGGCG   2880
GGCACCTGCG TCTCCCAGTA TGCCAACTGG CTGGTGGTGC TGCTCCTCGT CATCTTCCTG   2940
CTCGTGGCCA ACATCCTGCT GGTCAACTTG CTCATTGCCA TGTTCAGTTA CACATTCGGC   3000
AAAGTACAGG GCAACAGCGA TCTCTACTGG AAGGCGCAGC GTTACCGCCT CATCCGGGAA   3060
TTCCACTCTC GGCCCGCGCT GGCCCCGCCC TTTATCGTCA TCTCCCACTT GCGCCTCCTG   3120
CTCAGGCAAT TGTGCAGGCG ACCCCGGAGC CCCCAGCCGT CCTCCCCGGC CCTCGAGCAT   3180
TTCCGGGTTT ACCTTTCTAA GGAAGCCGAG CGGAAGCTGC TAACGTGGGA ATCGGTGCAT   3240
AAGGAGAACT TTCTGCTGGC ACGCGCTAGG GACAAGCGGG AGAGCGACTC CGAGCGTCTG   3300
AAGCGCACGT CCCAGAAGGT GGACTTGGCA CTGAAACAGC TGGGACACAT CCGCGAGTAC   3360
GAACAGCGCC TGAAAGTGCT GGAGCGGGAG GTCCAGCAGT GTAGCCGCGT CCTGGGGTGG   3420
GTGGCCGAGG CCCTGAGCCG CTCTGCCTTG CTGCCCCCAG GTGGGCCGCC ACCCCCTGAC   3480
CTGCCTGGGT CCAAAGACTG A
```

<u>SEQ ID NO:218 PAV9 Protein sequence:</u>
Protein Accession #:   none found

```
1          11         21         31         41         51
|          |          |          |          |          |
MEDAFGAAVV TVWDSDAHTT EKPTDAYGEL DFTGAGRKHS NFLRLSDRTD PAAVYSLVTR   60
TWGFRAPNLV VSVLGGSGGP VLQTWLQDLL RRGLVRAAQS TGAWIVTGGL HTGIGRHVGV   120
AVRDHQMAST GGTKVVAMGV AFWGVVRNRD TLINPKGSFP ARYRWRGDPE DGVQFPLDYN   180
YSAFFLVDDG THGCLGGENR FRLRLESYIS QQKTGVGGTG IDIPVLLLLI DGDEKMLTRI   240
ENATQAQLPC LLVAGSGGAA DCLAETLEDT LAPGSGGARQ GEARDRIRRF FPKGDLEVLQ   300
AQVERIMTRK ELLTVYSSED GSEEFETIVL KALVKACGSS EASAYLDELR LAVAWNRVDI   360
AQSELFRGDI QWRSFHLEAS LMDALLNDRP EFVRLLISHG LSLGHFLTPM RLAQLYSAAP   420
SNSLIRNLLD QASHSAGTKA PALKGGAAEL RPPDVGHVLR MLLGKMCAPR YPSGGAWDPH   480
PGQGFGESMY LLSDKATSPL SLDAGLGQAP WSDLLLWALL LNRAQMAMYF WEMGSNAVSS   540
ALGACLLLRV MARLEPDAEE AARRKDLAFK FEGMGVDLFG ECYRSSEVRA ARLLLRRCPL   600
WGDATCLQLA MQADARAFFA QDGVQSLLTQ KWWGDMASTT PIWALVLAFF CPPLIYTRLI   660
TFRKSEEEPT REELEFDMDS VINGEGPVGT ADPAEKTPLG VPRQSGRPGC CGGRCGGRRC   720
LRRWFHFWGA PVTIFMGNVV SYLLFLLLFS RVLLVDFQPA PPGSLELLLY FWAFTLLCEE   780
LRQGLSGGGG SLASGGPGPG HASLSQRLRL YLADSWNQCD LVALTCFLLG VGCRLTPGLY   840
HLGRTVLCID FMVFTVRLLH IFTVNKQLGP KIVIVSKMMK DVFFFLFFLG VWLVAYGVAT   900
EGLLRPRDSD FPSILRRVFY RPYLQIFGQI PQEDMDVALM EHSNCSSEPG FWAHPPGAQA   960
GTCVSQYANW LVVLLLVIFL LVANILLVNL LIAMFSYTFG KVQGNSDLYW KAQRYRLIRE   1020
FHSRPALAPP FIVISHLRLL LRQLCRRPRS PQPSSPALEH FRVYLSKEAE RKLLTWESVH   1080
KENFLLARAR DKRESDSERL KRTSQKVDLA LKQLGHIREY EQRLKVLERE VQQCSRVLGW   1140
VAEALSRSAL LPPGGPPPPD LPGSKD
```

SEQ ID NO:219 PBF1 DNA SEQUENCE
Nucleic Acid Accession #:   AA054237
Coding sequence: 1-894 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGGAGCCGC GGGCGCTCGT CACGCGCGCTC AGCCTCGGCC TCAGCCTGTG CTCCCTGGGG   60
CTGCTCGTCA CGGCCATCTT CACCGACCAC TGGTACGAGA CCGACCCCCG GCGCCACAAG   120
GAGAGCTGCG AGCGCAGCCG CGCGGGCGCC GACCCCCCGG ACCAGAAGAA CCGCCTGATG   180
```

```
ATCTTGGTAT TCAATGCTGA ATATGGAAAC AGCTCAGTTT TCTTGGAGAA CAGTACATTT    360
GATGAGTTTG GACATTCTAT CAATGATTAT TCAATATCTC CTGATGGGCA GTTTATTCTC    420
TTAGAATACA ACTACGTGAA GCAATGGAGG CATTCCTACA CAGCTTCATA TGACATTTAT    480
GATTTAAATA AAAGGCAGCT GATTACAGAA GAGAGGATTC CAAACAACAC ACAGTGGGTC    540
ACATGGTCAC CAGTGGGTCA TAAAATTGGCA TATGTTTGGA ACAATGACAT TTATGTTAAA    600
ATTGAACCAA ATTTACCAAG TTACAGAATC ACATGGACGG GGAAAGAAGA TATAATATAT    660
AATGGAATAA CTGACTGGGT TTATGAAGAG GAAGTCTTCA GTGCCTACTC TGCTCTGTGG    720
TGGTCTCCAA ACGGCACTTT TTTAGCATAT GCCCAATTTA ACGACACAGA AGTCCCACTT    780
ATTGAATACT CCTTCTACTC TGATGAGTCA CTGCAGTACC CAAAGACTGT ACGGGTTCCA    840
TATCCAAAGG CAGGAGCTGT GAATCCAACT GTAAAGTTCT TTGTTGTAAA TACAGACTCT    900
CTCAGCTCAG TCACCAATGC AACTTCCATA CAAATCACTG CTCCTGCTTC TATGTTGATA    960
GGGGATCACT ACTTGTGTGA TGTGACATGG GCAACACAAG AAAGAATTTC TTTGCAGTGG   1020
CTCAGGAGGA TTCAGAACTA TTCGGTCATG GATATTTGTG ACTATGATGA ATCCAGTGGA   1080
AGATGGAACT GCTTAGTGGC ACGGCAACAC ATTGAAATGA GTACTACTGG CTGGGTTGGA   1140
AGATTTAGGC CTTCAGAACC TCATTTTACC CTTGATGGTA ATAGCTTCTA CAAGATCATC   1200
AGCAATGAAG AAGGTTACAG ACACATTTGC TATTTCCAAA TAGATAAAAA AGACTGCACA   1260
TTTATTACAA AAGGCACCTG GGAAGTCATC GGGATAGAAG CTCTAACCAG TGATTATCTA   1320
TACTACATTA GTAATGAATA TAAAGGAATG CCAGGAGGAA GGAATCTTTA TAAAATCCAA   1380
CTTATTGACT ATACAAAAGT GACATGCCTC AGTTGTGAGC TGAATCCGGA AAGGTGTCAG   1440
TACTATTCTG TGTCATTCAG TAAAGAGGCG AAGTATTATC AGCTGAGATG TTCCGGTCCT   1500
GGTCTGCCCC TCTATACTCT ACACAGCAGC GTGAATGATA AAGGGCTGAG AGTCCTGGAA   1560
GACAATTCAG CTTTGGATAA AATGCTGCAG AATGTCCAGA TGCCCTCCAA AAAACTGGAC   1620
TTCATTATTT TGAATGAAAC AAAATTTTGG TATCAGATGA TCTTGCCTCC TCATTTTGAT   1680
AAATCCAAAA AATATCCTCT ACTATTAGAT GTGTATGCAG GCCCATGTAG TCAAAAAGCA   1740
GACACTGTCT TCAGACTGAA CTGGGCCACT TACCTTGCAA GCACAGAAAA CATTATAGTA   1800
GCTAGCTTTG ATGGCAGAGG AAGTGGTTAC CAAGGAGATA AGATCATGCA TGCAATCAAC   1860
AGAAGACTGG GAACATTTGA AGTTGAAGAT CAAATTGAAG CAGCCAGACA ATTTTCAAAA   1920
ATGGGATTTG TGGACAACAA ACGAATTGCA ATTTGGGGCT GGTCATATGG AGGGTACGTA   1980
ACCTCAATGG TCCTGGGATC GGGAAGTGGC GTGTTCAAGT GTGGAATAGC CGTTGCGCCT   2040
GTATCCCGGT GGGAGTACTA TGACTCAGTG TACACAGAAC GTTACATGGG TCTCCCAACT   2100
CCAGAAGACA ACCTTGACCA TTACAGAAAT TCAACAGTCA TGAGCAGAGC TGAAAATTTT   2160
AAACAAGTTG AGTACCTCCT TATTCATGGA ACAGCAGATG ATAACGTTCA CTTTCAGCAG   2220
TCAGCTCAGA TCTCCAAAGC CCTGGTCGAT GTTGGAGTGG ATTTCCAGGC AATGTGGTAT   2280
ACTGATGAAG ACCATGGAAT AGCTAGCAGC ACAGCACACC AACATATATA TACCCACATG   2340
AGCCACTTCA TAAAACAATG TTTCTCTTTA CCTTAGCACC TCAAAATACC ATGCCATTTA   2400
AAGCTTATTA AAACTCATTT TTGTTTTCAT TATCTCAAAA CTGCACTGTC AAGATGATGA   2460
TGATCTTTAA AATACACACT CAAATCAAGA AACTTAAGGT TACCTTTGTT CCCAAATTTC   2520
ATACCTATCA TCTTAAGTAG GGACTTCTGT CTTCACAACA GATTATTACC TTACAGAAGT   2580
TTGAATTATC CGGTCGGGTT TTATTGTTTA AAATCATTTC TGCATCAGCT GCTGAAACAA   2640
CAAATAGGAA TTGTTTTTAT GGAGGCTTTG CATAGATTCC CTGAGCAGGA TTTTAATCTT   2700
TTTCTAACTG GACTGGTTCA AATGTTGTTC TCTTCTTTAA AGGGATGGCA AGATGTGGGC   2760
AGTGATGTCA CTAGGGCAGG GACAGGGATAA GAGGGATTAG GGAGAGAAGA TAGCAGGGCA   2820
TGGCTGGGAA CCCAAGTCCA AGCATACCAA CACGAGCAGG CTACTGTCAG CTCCCCTCGG   2880
AGAAGAGCTG TTCACCACGA GACTGGCACA GTTTTCTGAG AAAGACTATT CAAACAGTCT   2940
CAGGAAATCA AATATCGAAA GCACTGACTT CTAAGTAAAC CACAGCAGTT GAAAGACTCC   3000
AAAGAAATGT AAGGGAAACT GCCAGCAACG CAGCCCCCAG GTGCCAGTTA TGGCTATAGG   3060
TGCTACAAAA ACACAGCAAG GGTGATGGGA AAGCATTGTA AATGTGCTTT TAAAAAAAAA   3120
TACTGATGTT CCTAGTGAAA GAGGCAGCTT GAAACTGAGA TGTGAACACA TCAGCTTGCA   3180
CTGTTAAAAG ATGAAAATAT TTGTATCACA AATCTTAACT TGAAGGAGTC CTTGCATCAA   3240
TTTTTCTTAT TTCATTTCTT TGAGTGTCTT AATTAAAAGA ATATTTTAAC TTCCTTGGAC   3300
TCATTTTAAA AAATGGAACA TAAAATACAA TGTTATGTAT TATTATTCCC ATTCTACATA   3360
CTATGGAATT TCTCCCAGTC ATTTAATAAA TGTGCCTTCA TTTTTTTC
```

Protein Accession #:   NP_001926.1

```
 1          11         21         31         41         51
 |          |          |          |          |          |
MKTPWKILLG LLGAAALVTI ITVPVVLLNK GTDDATADSR KTYTLTDYLK NTYRLKLYSL    60
RWISDHEYLY KQENNILVFN AEYGNSSVFL ENSTFDEFGH SINDYSISPD GQFILLEYNY   120
VKQWRHSYTA SYDIYDLNKR QLITEERIPN NTQWVTWSPV GHKLAYVWNN DIYVKIEPNL   180
PSYRITWTGK EDIIYNGITD WVYEEEVFSA YSALWWSPNG TFLAYAQFND TEVPLIEYSF   240
YSDESLQYPK TVRVPYPKAG AVNPTVKFFV VNTDSLSSVT NATSIQITAP ASMLIGDHYL   300
CDVTWATQER ISLQWLRRIQ NYSVMDICDY DESSGRWNCL VARQHIEMST TGWVGRFRPS   360
EPHFTLDGNS FYKIISNEEG YRHICYFQID KKDCTFITKG TWEVIGIEAL TSDYLYYISN   420
EYKGMPGGRN LYKIQLIDYT KVTCLSCELN PERCQYYSVS FSKEAKYYQL RCSGPGLPLY   480
TLHSSVNDKG LRVLEDNSAL DKMLQNVQMP SKKLDFIILN ETKFWYQMIL PPHFDKSKKY   540
PLLLDVYAGP CSQKADTVFR LNWATYLAST ENIIVASFDG RGSGYQGDKI MHAINRRLGT   600
FEVEDQIEAA RQFSKMGFVD NKRIAIWGWS YGGYVTSMVL GSGSGVFKCG IAVAPVSRWE   660
YYDSVYTERY MGLPTPEDNL DHYRNSTVMS RAENFKQVEY LLIHGTADDN VHFQQSAQIS   720
KALVDVGVDF QAMWYTDEDH GIASSTAHQH IYTHMSHFIK QCFSLP
```

Nucleic Acid Accession #:   none found
Coding sequence: 1-261 (underlined sequences correspond to start and stop codons)

```
 1          11         21         31         41         51
 |          |          |          |          |          |
```

```
ATGGCTCTGG CGAAGGTGAG GGAGCCAAAC GCAAATGACA ATGCCATCAG AGTTGACAAC    60
AGAAGTGTGA TTAAAGTGCG TGCTAACCAG TGTTCCCTGC ATGAGGCAGA AAGTGAATCC   120
AGAAACCCTC AGGAGCTCTG GATGGGCCTG CTCCTCTTGA TGGGGGTCCT AGAAGCATGT   180
GTGGAAATGA GGCCTCTGTC AGTCTGGTCC CTGAGAGATG ACAAGGAGCA GAGCCCCCAC   240
CAGCCCCACAC TGGATGTCTA A
```

```
1           11          21          31          41          51
|           |           |           |           |           |
MALAKVREPN ANDNAIRVDN RSVIKVRANQ CSLHEAESES RNPQELWMGL LLLMGVLEAC    60
VEMRPLSVWS LRDDKEQSPH QPTLDV
```

```
1           11          21          31          41          51
|           |           |           |           |           |
ATGCCAAATG CTGAGTTAGA AGCAAAGAGC CTTGGAAGCA GTAAATGTTT AAAAACTGCT    60
CTCATACTTG CTGTATGTTG TGGATCAGCA AATATAGTCA GCCCTCTACT TGAGCAAAAT   120
ATTGATGTAT CTTCTCAAGA TCTGGACAGA CGGCCAGAGA GTATGCTGTT TCTAGTCATC   180
ATCATGTGGA CCAGTTTTGT GGAAGACAAT CTTTCCATGG GCTGGGGGAA GCTAGAAGAT   240
TTTATGGCTA TTGAAGAAGA AATGAAGAAG CACGGAAGTA CTCATGTGGG ATTCCCAGAA   300
AACCTGACTA ATGGTGCCGC TGCTGGCAAT GGTGATGATG GATTAATTCC TCCAAGGAAG   360
AGCAGAACAC CTGAAAGCCA GCAATTTCCT GACACTGAGA ATGAAGAGTA TCACAGGTTT   420
GTCAAAGATC AGATAGTTGT AGATATGCGG CGTTATTTCT GA
```

```
1           11          21          31          41          51
|           |           |           |           |           |
MPNAELEAKS LGSSKCLKTA LILAVCCGSA NIVSPLLEQN IDVSSQDLDR RPESMLFLVI    60
IMWTSFVEDN LSMGWGKLED FMAIEEEMKK HGSTHVGFPE NLTNGAAAGN GDDGLIPPRK   120
SRTPESQQFP DTENEEYHRF VKDQIVVDMR RYF
```

```
1           11          21          31          41          51
|           |           |           |           |           |
GACTGCTTGC ATTAAAGGAC TTCCTCATCC TTTTTTTCAT GAAACTGAGC TTGCTTAATC    60
AGAGATGGAG CAAACTGACT GCAAACCCTA CCAGCCTCTA CCAAAAGTCA AGCATGAAAT   120
GGATCTAGCT TACACCAGTT CTTCTGATGA GAGTGAAGAT GGAAGAAAAC CAAGACAGTC   180
ATACAACTCC AGGGAGACCC TGCACGAGTA TAACCAGGAG CTGAGGATGA ATTACAATAG   240
CCAGAGTAGA AAGAGGAAAG AAGTAGAAAA ATCTACTCAA GAGATGGAAT TCTGTGAAAC   300
CTCTCACACT CTGTGCTCTG GCTACCAAAC AGACATGCAC AGCGTTTCTC GGCATGGCTA   360
CCAGCTAGAG ATGGGATCTG ATGTGGACAC AGAGACAGAA GGTGCTGCCT CACCTGACCA   420
TGCACTAAGA ATGTGGATAA GGGGAATGAA ATCAGAGCAT AGTTCCTGTT TGTCCAGCCG   480
GGCCAACTCT GCATTATCCT TGACTGACAC TGACCATGAA AGGAAGTCTG ATGGGGAAAA   540
TGGTTTCAAA TTCTCTCCTG TTTGTTGTGA CATGGAGGCT CAAGCTGGGT CTACTCAAGA   600
TGTGCAGAGC AGCCCACACA ACCAGTTCAC CTTCAGACCC CTCCCACCGC CACCTCCGCC   660
TCCTCATGCC TGCACCTGTG CCAGGAAGCC ACCCCCTGCA GCGGACTCTC TTCAGAGGAG   720
ATCAATGACT ACCCGCAGCC AGCCCAGCCC AGCTGCTCCA GCTCCCCCAA CCAGCACGCA   780
GGATTCAGTC CATCTGCATA ACAGCTGGGT CCTGAACAGC AACATACCAT TGGAGACCAG   840
GCATTCCCTG TTCAAACATG GATCTGGTTC CTCTGCGATC TTCAGTGCAG CCAGTCAGAA   900
CTACCCTCTG ACATCCAATA CCGTGTACTC GCCCCCTCCC AGGCCTCTTC CTCGAAGCAC   960
CTTTTCCCGA CCTGCCTTTA CCTTTAACAA ACCTTACAGG TGCTGCAACT GGAAGTGCAC  1020
AGCATTGAGC GCCACTGCAA TCACAGTGAC TTTGGCCTTG TTACTAGCCT ATGTGATTGC  1080
AGTGCATTTG TTCGGCCTGA CTTGGCAGTT GCAACCAGTT GAAGGAGAGC TGTATGCAAA  1140
TGGAGTTAGC AAAGGGAACA GGGGGACCGA GTCCATGGAC ACTACTTACT CTCCAATTGG  1200
AGGAAAAGTT TCTGATAAAT CAGAGAAAAA AGTGTTTCAG AAGGGACGGG CGATAGACAC  1260
TGGAGAAGTT GACATTGGTG CACAGGTCAT GCAGACCATT CCACCTGGTT TATTCTGGCG  1320
TTTCCAGATT ACTATCCACC ATCCAATATA TCTGAAGTTC AATATTTCTT TAGCCAAGGA  1380
CTCTCTGCTG GGAATTTATG GCAGAAGAAA CATTCCACCT ACACATACTC AGTTTGATTT  1440
TGTAAAACTA ATGGATGGCA AACAGCTGGT CAAGCAGGAC TCCAAGGGCC TGATGATAC   1500
ACAGCACTCC CCTCGGAACC TGATCTTAAC TTCGCTTCAG GAGACAGGTT TCATAGAGTA  1560
TATGGATCAA GGACCTTGGT ATCTGGCGTT TTACAATGAT GGAAAAAAGA TGGGAGCAAGT  1620
ATTCGTGTTA ACTACAGCAA TTGAAATAAT GGATGACTGT TCAACCAATT GCAATGGAAA  1680
TGGAGAGTGT ATCTCTGGCC ATTGTCATTG TTTCCCAGGA TTCCTTGGAC CTGACTGTGC  1740
TAGAGATTCC TGCCCTGTGC TGTGTGGTGG GAATGGAGAA TACGAGAAAG GACACTGTGT  1800
CTGCCGGCAT GGCTGGAAGG GGCCAGAGTG TGACGTTCCG GAAGAACAAT GCATTGATCC  1860
AACATGCTTT GGGCCACGGCA CCTGCATCAT GGGAGTCTGC ATCTGTGTCA CAGGATACAA  1920
```

```
        AGGAGAAATA TGCGAGGAAG AGGACTGCCT AGACCCAATG TGTTCCAACC ATGGCATCTG  1980
        TGTAAAAGGA GAATGTCACT GTTCTACTGG CTGGGGAGGA GTTAACTGTG AAACACCACT  2040
        TCCTGTATGT CAAGAGCAGT GCTCAGGACA CGGAACTTTT CTTCTGGACG CTGGAGTATG  2100
   5    CAGCTGTGAT CCCAAGTGGA CAGGATCTGA CTGCTCAACA GAGCTGTGTA CCATGGAGTG  2160
        TGGTAGCCAT GGAGTCTGCT CAAGAGGAAT TTGCCAGTGT GAAGAAGGCT GGGTAGGACC  2220
        AACATGTGAG GAACGCTCCT GTCATTCTCA TTGTACTGAG CATGGCCAAT GCAAAGATGG  2280
        AAAATGTGAG TGTAGCCCTG GATGGGAGGG CGACCACTGC ACAATTGCTC ACTACTTAGA  2340
        TGCTGTCCGA GATGGCTGCC CAGGGCTCTG CTTTGGAAAT GGACGATGTA CCCTGGATCA  2400
  10    AAATGGTTGG CACTGTGTGT GTCAGGTGGG TTGGAGTGGG ACAGGCTGCA ATGTTGTCAT  2460
        GGAAATGCTT TGTGGAGATA ACTTGGACAA TGATGGAGAT GGTTTAACCG ACTGTGTGGA  2520
        TCCTGACTGT TGTCAACAAA GCAACTGTTA TATAAGTCCT CTCTGCCAGG GCTCACCAGA  2580
        TCCTCTTGAC CTCATTCAGC AAAGCCAAAC TCTCTTCTCT CAGCACACTT CAAGACTTTT  2640
        TTATGATCGA ATCAAATTCC TCATTGGCAA GGACAGTACT CATGTCATTC CTCCTGAGGT  2700
  15    GTCATTTGAC AGCAGGCGTG CCTGTGTGAT TCGAGGCCAA GTGGTGGCCA TAGATGGAAC  2760
        TCCTCTAGTG GGAGTGAATG TCAGTTTCTT GCACCACAGT GATTATGGGT TTACCATCAG  2820
        CCGGCAAGAT GGAAGCTTTG ACCTCGTGGC CATCGGTGGC ATCTCTGTCA TCTTAATCTT  2880
        CGACCGATCC CCTTTCCTGC CTGAGAAGAG AACACTCTGG TTGCCTTGGA ATCAGTTTAT  2940
        TGTGGTAGAG AAAGTCACCA TGCAGAGAGT TGTATCAGAC CCGCCATCCT GCGATATCTC  3000
  20    CAACTTTATC AGCCCAAACC CTATTGTGCT TCCTTCACCG CTCACATCAT TTGGAGGGTC  3060
        CTGTCCAGAG AGGGGAACTA TTGTTCCTGA GCTGCAGGTT GTACAGGAGG AAAATTCCCAT  3120
        TCCCTCCAGC TTTGTGAGGC TGAGTTACCT GAGCAGCCGC ACCCCTGGGT ATAAAACCCT  3180
        GCTACGGATC CTTCTGACAC ATTCAACGAT TCCCGTAGGC ATGATAAAAG TACACCTCAC  3240
        AGTAGCTGTG GAAGGGCGAC TCACACAGAA GTGGTTTCCC GCCGCAATTA ATCTTGTCTA  3300
  25    CACATTTGCT TGGAACAAGA CCGATATCTA TGGACAGAAG GTTTGGGGCC TGGCAGAGGC  3360
        TTTGGTATCT GTGGGGATATG AATATGAAAC GTGCCCTGAC TTTATTCTCT GGGAGCAAAG  3420
        GACAGTCGTT TTACAAGGTT TTGAGATGGA TGCTTCTAAC CTAGGAGACT GGTCTTTGAA  3480
        TAAGCATCAC ATTTTGAATC CTCAAAGTGG AATCATACAT AAAGGGAATG GAGAAAATAT  3540
        GTTCATTTCC CAGCAGCCCC CAGTCATATC AACCATAATG GGTAATGGAC ACCAAAGGAG  3600
  30    TGTAGCCTGC ACCAACTGCA ATGGCCCAGC CCACAACAAC AAACTCTTTG CTCCTGTCGC  3660
        CTTAGCTTCT GGCCCTGATG GCAGTGTGTA TGTTGGCGAC TTCAATTTTG TAAGGAGAAT  3720
        ATTTCCCTCG GGAAACTCCG TTAGTATTTT GGAATTAAGC ACAAGTCCTG CTCACAAATA  3780
        CTATCTGGCT ATGGACCCTG TGTCTGAATC ACTCTATCTA TCAGACACCA ATACTCGCAA  3840
        AGTCTACAAG TTGAAATCTC TTGTGGAGAC GAAAGATCTG TCCAAGAATT TTGAAGTGGT  3900
  35    GGCAGGAACT GGTGATCAGT GCCTTCCCTT TGACCAGAGT CATTGTGGAG ATGGTGGGAG  3960
        AGCATCGGAA GCTTCACTGA ATAGCCCTCG AGGCATCACA GTTGATAGGC ATGGATTTAT  4020
        TTACTTTGTG GATGGGACTA TGATTCGCAA AATTGATGAG AATGCTGTGA TCACAACTGT  4080
        AATCGGCTCA AATGGTCTGA CTTCCACACA ACCACTGAGC TGTGACTCAG GAATGGACAT  4140
        CACTCAGGTG CGATTAGAGT GGCCAACAGA CCTTGCAGTA AATCCTATGG ACAATTCATT  4200
  40    GTATGTCTTG GATAACAACA TTGTGCTGCA AATTTCTGAC AACAGGCGTG TTCGGGATCAT  4260
        CGCAGGACGC CCCATTCACT GCCAGGTGCC AGGCATCGAT CATTTCCTGG TCAGCAAGGT  4320
        AGCAATTCAC TCCACTCTAG AGTCAGCGAG GGCCATCAGT CTTCCCACA GCGGGCTGCT  4380
        CTTCATAGCT GAAACAGACG AGAGGAAAGT AAACCGCATT CAGCAAGTAA CCACCAATGG  4440
        GGAGATCTAC ATCATCGCTG GTGCCCCCAC TGACTGTGAC TGCAAAATTG ATCCAAACTG  4500
  45    TGACTGTTTT TCAGGTGATG GTGGCTATGC CAAAGATGCA AAGATGAAAG CCCCTTCCTC  4560
        CTTAGCAGTG TCGCCTGATG GAACCCTCTA TGTGGCAGAC CTCGGAAATG TTCGAATTCG  4620
        TACCATCAGC AGGAACCAAG CCCACCTGAA TGACATGAAC ATTTATGAGA TTGCTTCACC  4680
        CGCTGATCAG GAACTGTACC AGTTCACTGT AAATGGAACC CACCTACACA CCCTGAACTT  4740
        GATAACAAGG GACTATGTTT ATAACTTCAC CTACAATTCT GAAGGTGACT TGGGCGCGAT  4800
  50    TACCAGCAGC AATGGCAATT CAGTGCACAT TCGCCGTGAT GCAGGCGGAA TGCCGCTATG  4860
        GCTTGTGGTG CCTGGCGGAC AAGTATACTG GCTGACTATA AGCAGCAATG GAGTCCTGAA  4920
        AAGAGTGTCA GCCCAAGGCT ATAATCCGGC CTTAATGACC TATCCAGGAA ACACAGGGCT  4980
        TCTGGCTACC AAAAGTAACG AAAATGGATG GACAACCGTT TATGAGTATG ACCCCGAGGG  5040
        ACACCTGACC AATGCAACGT TTCCCACTGG AGAGGTCAGC AGCTTCCACA GTGACCTGGA  5100
  55    GAAGCTGACA AAAGTGGGAGC TAGATACTTC CAACCGTGAA AATGTCCTCA TGTCAACCAA  5160
        CTTGACGGCA ACTAGTACCA TATATATTTT AAAACAAGAA AATACTCAAA GTACCTATCG  5220
        GGTGAATCCA GATGGTTCCC TGCGTGTCAC TTTTGCCAGC GGGATGGAGA TCGGCCTCAG  5280
        CTCAGAGCCC CACATCCTGG CAGGGGCAGT CAACCCTACC CTGGGCAAAT GCAACATCTC  5340
        ATTGCCCGGA GAGCACAATG CAAACCTCAT CGAGTGGCGG CAGAGGAAGG AGCAAAACAA  5400
  60    AGGCAATGTT TCGGCTTTTG AAAGGAGGCT GAGGGCCCGAC AACAGAAACC TACTCTCCAT  5460
        AGATTTTGAT CATATAACCC GCACAGGAAA GATCTATGAT GACCATCGAA AATTCACCCT  5520
        TCGAATTCTT TATGACCAGA CTGGGCGACC CATTCTGTGG TCTCCTGTAA GCAGATATAA  5580
        TGAAGTGAAC ATCACATATT CACCTTCGGG ATTGGTGACG TTTATTCAAA GAGGAACGTG  5640
        GAATGAAAAA ATGGAATATG ACCAGAGTGG GAAAATTATT TCAAGAACTT GGGCTGATGG  5700
  65    GAAAATTTGG AGCTATACCT ACTTAGAAAA ATCTGTGATG CTTCTCCTAC ACAGCCAGCG  5760
        GCGTTACATC TTTGAGTATG ACCAATCAGA TTGCCTGCTG TCAGTTACCA TGCCTAGCAT  5820
        GGTGCGCCAC AGCTTACAAA CCATGCTTTC AGTGGGCTAC TACCGTAATA TCTACACCCC  5880
        ACCGGACAGT AGCACTTCTT TTATCCAAGA CTATAGTCGA GATGGCCGAT TGCTACAGAC  5940
        CCTGCATCTG GGGACAGGGC GCAGAGTCTT ATACAAGTAC ACCAAGCAAG CAAGGCTTTC  6000
  70    TGAGGTTCTC TATGATACCA CTCAGGTCAC ATTAACATAT GAAGAGTCTT CTGGAGTGAT  6060
        TAAGACAATA CACCTGATGC ATGACGGATT CATCTGCACA ATCAGATACA GGCAAACAGG  6120
        ACCTCTTATT GGACGCCAGA TTTTCAGATT CAGTGAAGAA GGCCTTGTGA ATGCACGGTT  6180
        CGACTACAGC TACAACAATT TCCGAGTCAC AAGCATGCAA GCTGTAATCA ATGAAACCCC  6240
        TTTGCCTATA GATCTTTACC GATATGTTGA TGTCTCTGGC AGAACAGAGC AGTTTGGAAA  6300
  75    ATTCAGTGTA ATTAATTACG ATTTAAATCA GGTCATAACT ACTACAGTGA TGAAACACAC  6360
        CAAAATCTTC AGTGCCAATG GACAAGTCAT TGAAGTCCAA TATGAAATCC TAAAGGCAAT  6420
        TGCCTACTGG ATGACCATTC AATATGATAA TGTGGGCCGA CATGGTAATA TGTGCATAAG  6480
        GGTAGGAGTA GATGCCAATA TAACAAGGTA CTTCTATGAA TACGATGCTG ATGGGCAACT  6540
        TCAGACTGTT TCTGTAAATG ACAAAACCCA GTGGCGTTAT AGTTACGATC TGAATGGAGA  6600
  80    CATCAACCTC TTAAGCCATG GGAAGAGTGC TCGTCTTACT CCTCTCCGAT ATGACCTCCG  6660
        AGACCGCATC ACCAGATTAG GAGAAATTCA GTATAAAATG GATGAAGATG GCTTTCTGAG  6720
```

395

```
     GCAGAGGGGA AATGATATTT TTGAATATAA TTCTAATGGC CTGCTGCAGA AAGCCTACAA  6780
     TAAGGCTTCT GGCTGGACTG TGCAGTATTA CTATGATGGG CTTGGGCGAC GTGTCGCGAG  6840
     TAAGTCCAGC CTAGGGCAGC ACCTTCAGTT CTTTGTCGAC GCGACCGCGA ACCCCATAAG  6900
     AGTTACTCAT TTGTACAACC ACACAAGCTC GGAGATTACA TCTCTGTATT ATGATCTCCA  6960
  5  AGGTCACCTT ATTGCCATGG AGTTAAGCAG TGGTGAAGAA TATTATGTAG CCTGTGATAA  7020
     TACAGGTACC CCACTAGCTG TGTTCAGCAG CCGAGGTCAG GTCATAAAGG AGATACTATA  7080
     CACACCTTAT GGCGATATCT ATCATGACAC TTACCCTGAC TTTCAGGTCA TAATTGGTTT  7140
     TCATGGAGGA CTCTATGATT TCCTTACTAA ATTAGTGCAC CTGGGGCAAA GGGATTATGA  7200
     TGTTGTTGCT GGCAGATGGA CAACGGCCTA TCATCACATA TGGAAACAGT TGAACCTCCT  7260
 10  TCCTAAACCA TTCAACCTCT ACTCCTTTGA AAATAACTAC CCAGTTGGCA AAATTCAAGA  7320
     TGTTGCAAAG TATACCACAG ACATCAGAAG TTGGTTGGAG CTATTTGGTT TCCAATTACA  7380
     CAATGTACTA CCTGGATTTC CCAAACCTGA ATTAGAAAAT TTAGAATTAA CTTACGAGCT  7440
     TCTACGGCTT CAGACAAAAA CTCAAGAGTG GGATCCTGGA AAGACTATCC TGGGCATTCA  7500
     GTGTGAACTC CAGAAACAGC TCAGGAATTT CATTTCCTTG GACCAACTAC CTATGACTCC  7560
 15  CCGATACAAT GATGGACGGT GCCTTGAAGG AGGGAAGCAA CCAAGGTTTG CTGCTGTCCC  7620
     TTCTGTTTTT GGGAAAGGTA TAAAATTTGC CATCAAGGAT GGCATAGTAA CAGCTGATAT  7680
     TATAGGAGTA GCCAATGAAG ATAGCAGGCG GCTTGCTGCC ATTCTCAATA ATGCCCATTA  7740
     CCTGGAAAAC CTACATTTTA CCATAGAGGG GAGGGACACT CACTACTTCA TTAAGCTTGG  7800
     GTCTCTGGAG GAAGACCTGG TGCTCATCGG TAACACTGGG GGGAGGCGGA TTCTGGAGAA  7860
 20  TGGTGTCAAT GTCACTGTGT CCCAGATGAC TTCTCTGTTG AATGGGAGGA CTAGACGGTT  7920
     TGCAGATATT CAGCTCCAGC ATGGAGCCCT GTGCTTCAAC ATCCGGTATG GGACAACTGT  7980
     CGAAGAGGAA AAGAATCACG TGTTGGAGAT TGCCAGACAG CGCGCAGTGG CCCAGGCCTG  8040
     GACTAAGGAA CAAAGAAGGC TGCAAGAGGG GGAAGAGGGG ATTAGGGCAT GGACAGAAGG  8100
     GGAAAAGCAG CAGCTTTTGA GCACTGGGCG GGTACAAGGT TACGATGGGT ATTTTGTTTT  8160
 25  GTCTGTTGAG CAGTATTTAG AACTTTCTGA CAGTGCCAAT AATATTCACT TTATGAGACA  8220
     GAGCGAAATA GGCAGGAGGT AACAAAAATA TCTCTGCCTT TGCGTCACCA AAGACTGCCT  8280
     GTTTTTAAAA CATAAAATGG TTTATTGTAT TGGTTTTCTA GATCAGAACT CTGTATATGT  8340
     AAATATGGAG GAAAAACATA TCCAACTGCC TTTCAATGTG ACGGAAGATG GTATTTTAAT  8400
     ATTGTTTGTT TAAACTCTTT AAGAAATGAC AGAGATTTTT AGTTCTTGTG TGGCAGTATT  8460
 30  CAAAATAACA CAAGTAGAAC TCAAACAGCT AAAAACAGTT TTCAGAAAGC ACCACTTTCA  8520
     ATTTGCCGAG CCATGCATAT GTTCCAATAT CCAGAAAGAA CCCAAGGTTC TCTATCTCTA  8580
     TTGTGAGAAG CAGTTTCATC CTTAACTGTT GGCAGAACTT ACGGGCTATT TGAATAGGTG  8640
     GTGCAATAGT ATCTGAAACT TGCCTTTCGA AAGACTGCCA GCCCTTTGAC GTTTTCCAGA  8700
     TCTGTTATAG GAAACTTAAA AACAGGTGTA AAATGTCTTC AGCCACCATC TCCTAGAGTG  8760
 35  AGGACCCAAT TGCCCTTCCT TCTTGATTAT TCCTCCTTGC TTGTTAAAGT AAATGCCATA  8820
     TTGTTGTGCT GTGTTTTGGC GTGTGGTGGC TGGGTTCTGT CTACCATGCT TCCCTGTGGG  8880
     TGTGGTAACC AGACTGTATA GCCGCTATTT GCTCGTGTGT ACATGATACC AAAGCAGCTG  8940
     GCCAGCGTGA CCTCTCTCAC ACGACCTGTT TTGACTCAAT TTTTTACTAA AAGTTGTTCA  9000
     GCTGTATTGG TATCATGTAA ACATAGCTTT TATTAACCTG GGTAGGAATT TCTCATTTAT  9060
 40  ATATAGGATG TGTTTTGGTC ATAGTTTCAC ATTAGTGATT CAGTATCTAT ACACTGACCC  9120
     AATGGTTTTG TGCACATGAA CGGTAATTTA CTTAAAAGTA TGATTCTGGT ACAAAAACAA  9180
     ACAAAGGCTT TAGCAGGCAT ACGTGTCTGG GATGCCGATA CATACATTAA CTACTACTGC  9240
     AGAAATTCAT AAGAGCCAAA ACCTTAAAAA AATAGACCTG GTACTTAAGT GAAAGTACTA  9300
     AAGGGAAGAC CAGACCAAAC ATCACAGCAG TTGCTGCCAC ATTGTTTCAG CCCACTTAGA  9360
 45  TTTATCTTTC AAATGTACAA TTCTGTATTG AACATCTCCC AGCCATCTTC AGGAAATCGA  9420
     ATCAAGTAAA TCCTTTCCAA CCGAAAACAT TTCAACTAAC TATAGAGAGG CAGACTCATT  9480
     TTTATTAAAA TAATTTATAC AGTTAGTTAT TTTCGTTCTC CGTACTTACC CATTTATCTT  9540
     TATTTAATCG TCTCTACTGC CTAGGAAAAT AACTATTTTC CAGGACGGGT TATTTGTTCT  9600
     GCGATCATTT AAAATTTGGA GAAAGGTCAG GATTAGTGTT AATATCAGCT GCAGTTTCTC  9660
 50  AATCTCTAGG AATCCTGCAG TAAAACAAGC CCCTTGGTGA GCTGGAAGAT TTGTGCCCAG  9720
     TGACAAAGAG ATAGTTTGTA AAATGCTGTG TAATTGTAAG TTACCACAAA TGAAAATACA  9780
     TGACAGCACA ATGTGGCCCG TAGAAAATTC CCCTGAGCCA GCTTCTGCAC TTTCATCACC  9840
     GAATCTGAAC ATTTGCTATG TCTGAAGGCA AATTTATGAT GGAATGTTAG TTTGGATTCT  9900
     TTCCAGATGC TACCTAAATG CAGTGTGGGG TCATTGCCTT GCTTTGCGAT GACAGTTTCT  9960
 55  TTGAAAATAT GCAAAGTCAT AAGCTCATGT TAAGGTTTTT CAAGAGTCTG CCTCCTACTA 10020
     CACAAAGGAA AGCAAGGGAA AGGAAATGAC CCTGGCAAAC AGTAGGGAAG GGTGTATTCA 10080
     AACATTTCAT TTTCAAAACC TTCGGGTTAG AATACCACTT ACACATGTAT TCTGAGAGAC 10140
     AGAATTCATG AGGAACTCAT CTCTCTTTAT AACTGGAAAC ACACCAGCTT GATATATTGC 10200
     TAATCCATAC TAAAATCATA TTATTGGGTT TTTTCTGAAT CAGGCCTGTA TTAATGGTAC 10260
 60  AGTATTTATT CAGAATGGAA TTCTAAAATT ACTAACAAAC TTGTTGAAAA TTTGAATACC 10320
     TCCACACCAA CCTAAAAATG GACCTTAAGT TCCTAGAACC TCTGATGTTC TTTTAAATTA 10380
     ATGGAAAAAT AATTTGTGAA CTGTATATAG AGAGTGCATT CATAAATGTG ATTATGTATT 10440
     TTATCACAAA TCCAAAATGT CAATATTAGA GTCTATTTTG CTTATATTTT AAGCAATTAT 10500
     ACGTTTTTGC AATTCATTGA TGATGTATCA TTTTCAAACT GCTTTAAATA TCCATTAGAA 10560
 65  ACAAATATTT GAAGCTTTTA CTTAATAGTG ATTACCTTGA ACTGTGCATT TCTAGTTTGT 10620
     AATACGTATT TGGTTGGTTC GTGCCTTTAG TTTGTTAAAG TTACATTTGT ATTATATTCA 10680
     GGAAATGCAC TTTTTATTAC TTACAGCTGT GGTTTTAATA CTGCCTTGAA CTATTATTAT 10740
     TCTTTTTACA ACTCCTAAAG CTTGAGGGAG GAAAGAAAAA AAAAACAAAA CTACTAATCA 10800
     GTAGTAAATC GAAGAGAAAC ATTTTGGCAT TTCTTAAGAA GAAGATGGAG ATATTGAGTA 10860
 70  TATCACTTCC TATTCAGCTG AATAGAAAGA ATGCCTTCAT TGACTTGCAG TTCTGCAGTT 10920
     TAAATTATTG AAAGAACAAT TCGTTTGCAT TTCCTGATGA AAGTAAAAGC ATTTTTCAGA 10980
     GAAACATATG AATTTCTCAT ACCCAGCAGA CAGATGGCTG ACACTGCACA GCCACACACC 11040
     ATTCGAGTAA GTTAAAGTGA GAGCATAGTA GTTGGACTCT CCTATGAAGA ACATTCTGGG 11100
     CTGGAGGCAG GGAATACTCC ATGGTTGTTT CTTTTTCCTA CTTAAGCCCA TTTTGTTTGT 11160
 75  GCTTTTCTGT TTTGTTTTGT TTTCACTCTT GCACTACAGT CTAGAGATCC AAATGAACTG 11220
     AAAAGTTCAA AGTTTAACAC ATTTAAATAT GTTTACTTTT AGTTGTCATT CTAATCGTTA 11280
     TTGATTAGAA GCATGACTCC TGAAGGAAAG GGAAATAAAT CTCAATTCAT ACTAACTTGC 11340
     AACAAAACAC TTTTACCATA TAAATAAGTA TATGATTTAT TTTTAACCCA AAAAATGTAT 11400
     AAAATAAGTG TGTCCTTTAC TGTCAATTTA TCGAAGAGAT CTATAATATA TAGACTACAT 11460
 80  ATATATAATA TATACAACAT AGCCAAATGT ATGAAAACTT GACAATGTAT AATTTGGAAT 11520
```

396

```
TCACATGCTA CCTATGTAGA CAGGTATGAA ATTAAGTTAT AATTTTCATG AGACATTTTC 11580
ATCACTGTTG ACACAGTTTC AAGGCATTCC ATCATGTTAT TTTGACTCTT TTTCTTTTTT 11640
TTTTCTTTAA AAATATATTT TTAACTAGAC CAGGCCCCAC TATAATATCA CTTAAGAGAG 11700
TCAGGGCAAA GTTTTTGCAT TTATGAAGAT GTGTTCATGT AAGGGTGATT GTAATGGAGT 11760
TCATTGGTAA TAGAAGCAAA AGTACAGTAA CGAAGTATTG AAAAGAAAAT TTTGGAGACA 11820
TTGGAGCATA TTATATATAG CTTGTGGAAA GACATAAGGC TACAGATGGA ATGGAACATT 11880
CCTGTTTTCT TGAAGAAATT CACATACACA TAGCTGACCT GACTAGTACT TCAGCTCTTC 11940
CACAGCCTTC TATAAAGGTT CTTTCTTCTG CAAAGAAAAC AAAACAAAAC AAAACAAAAC 12000
AAAAAAAAAC AAAAAAAGCG CAAAAAACAA AAAAACAAAA AAAAGCAAAG TAAAATTTAA 12060
AAATACAGAA AACAAACAAC AAAAAAGAAT TCAACCATAA ATAGTGACTA TTATTTTCAG 12120
TGTGTCCTTC ATGTGAAAGC TATTAAGGAC CAAATATACT ACTGTTCATA AGAAGAAATT 12180
ACTTTCTAAA CAGTAACTGA AAATACTTAG AGTTAAACTT GCTGTGGATT TTGTCTTGGC 12240
AGTTGTCATC TTACATTATT TGTCAAAGGA AATGTGTTTG GCAGTTAAAA ATCTTTCCTT 12300
AGATTTAGTG GTGGACTTTA ACCTCTTAAA TAAATGTTAG TATATCAGAT TGTGTCCTTG 12360
AAAAATATTT TACTTGTATG AATCATGACA ACGTCTAAAT CTTTACTATT CTTCTGGCAA 12420
AAGCATCAGT AAGAAAGAAG GCGAAAAAGA GAAGTATAGC CTTTATGTCA GAAAAACATT 12480
CTTTTTAGCT GCTTACTTTC TCATGAAAAG TAAAGATGTT TACAGTGTAT GCCAAGTTTT 12540
CAGTTTCTGT ATAACAACAG GTAGAGGTTC TAATCATATT GAAAATTGTG TTATAATGGT 12600
CTGAGCCATG TTGCTAGGAA ACAATAGGTT CCAATTTTGT ATTCCTCTC TCCTGTGCTG 12660
AAAAGTGACT GGATACTGTA CAGGTTCATG TTCTCTGGCT GCAGTTAAAT GGTCTTTTGC 12720
ATTTTGCTCT GGCTTTCAGG CCAGAAGCAT GCATTTTTCT ACAAGAGCAT CACAACAACA 12780
TGCTGTAAAT ATTTAAAGTT AAACATTATG TGTTGATATT TGAAAGAAAA GTACTTTGAA 12840
TATTTCATTT TTAAAAAATA AAATTGCCAA TGAAAAAAAA
```

SEQ ID NO:230 PEZ2 Protein sequence:
Protein Accession #:   NP_055068

```
1          11         21         31         41         51
|          |          |          |          |          |
MEQTDCKPYQ PLPKVKHEMD LAYTSSSDES EDGRKPRQSY NSRETLHEYN QELRMNYNSQ  60
SRKRKEVEKS TQEMEFCETS HTLCSGYQTD MHSVSRHGYQ LEMGSDVDTE TEGAASPDHA  120
LRMWIRGMKS EHSSCLSSRA NSALSLTDTD HERKSDGENG FKFSPVCCDM EAQAGSTQDV  180
QSSPHNQFTF RPLPPPPPPP HACTCARKPP PAADSLQRRS MTTRSQPSPA APAPPTSTQD  240
SVHLHNSWVL NSNIPLETRH SLFKHGSGSS AIFSAASQNY PLTSNTVYSP PPRPLPRSTF  300
SRPAFTFNKP YRCCNWKCTA LSATAITVTL ALLLAYVIAV HLFGLTWQLQ PVEGELYANG  360
VSKGNRGTES MDTTYSPIGG KVSDKSEKKV FQKGRAIDTG EVDIGAQVMQ TIPPGLFWRF  420
QITIHHPIYL KFNISLAKDS LLGIYGRRNI PPTHTQFDFV KLMDGKQLVK QDSKGSDDTQ  480
HSPRNLILTS LQETGFIEYM DQGPWYLAFY NDGKKMEQVF VLTTAIEIMD DCSTNCNGNG  540
ECISGHCHCF PGFLGPDCAR DSCPVLCGGN GEYEKGHCVC RHGWKGPECD VPEEQCIDPT  600
CFGHGTCIMG VCICVPGYKG EICEEEDCLD PMCSNHGICV KGECHCSTGW GGVNCETPLP  660
VCQEQCSGHG TFLLDAGVCS CDPKWTGSDC STELCTMECG SHGVCSRGIC QCEEGWVGPT  720
CEERSCHSHC TEHGQCKDGK CECSPGWEGD HCTIAHYLDA VRDGCPGLCF GNGRCTLDQN  780
GWHCVCQVGW SGTGCNVVME MLCGDNLDND GDGLTDCVDP DCCQQSNCYI SPLCQGSPDP  840
LDLIQQSQTL FSQHTSRLFY DRIKFLIGKD STHVIPPEVS FDSRRACVIR GQVVAIDGTP  900
LVGVNVSFLH HSDYGFTISR QDGSFDLVAI GGISVILIFD RSPFLPEKRT LWLPWNQFIV  960
VEKVTMQRVV SDPPSCDISN FISPNPIVLP SPLTSFGGSC PERGTIVPEL QVVQEEIPIP  1020
SSFVRLSYLS SRTPGYKTLL RILLTHSTIP VGMIKVHLTV AVEGRLTQKW FPAAINLVYT  1080
FAWNKTDIYG QKVWGLAEAL VSVGYEYETC PDFILWEQRT VVLQGFEMDA SNLGDWSLNK  1140
HHILNPQSGI IHKGNGENMF ISQQPPVIST IMGNGHQRSV ACTNCNGPAH NNKLFAPVAL  1200
ASGPDGSVYV GDFNFVRRIF PSGNSVSILE LSTSPAHKYY LAMDPVSESL YLSDTNTRKV  1260
YKLKSLVETK DLSKNFEVVA GTGDQCLPFD QSHCGDGGRA SEASLNSPRG ITVDRHGFIY  1320
FVDGTMIRKI DENAVITTVI GSNGLTSTQP LSCDSGMDIT QVRLEWPTDL AVNPMDNSLY  1380
VLDNNIVLQI SENRRVRIIA GRPIHCQVPG IDHFLVSKVA IHSTLESARA ISVSHSGLLF  1440
IAETDERKVN RIQQVTTNGE IYIIAGAPTD CDCKIDPNCD CFSGDGGYAK DAKMKAPSSL  1500
AVSPDGTLYV ADLGNVRIRT ISRNQAHLND MNIYEIASPA DQELYQFTVN GTHLHTLNLI  1560
TRDYVYNFTY NSEGDLGAIT SSNGNSVHIR RDAGGMPLWL VVPGGQVYWL TISSNGVLKR  1620
VSAQGYNPAL MTYPGNTGLL ATKSNENGWT TVYEYDPEGH LTNATFPTGE VSSFHSDLEK  1680
LTKVELDTSN RENVLMSTNL TATSTIYILK QENTQSTYRV NPDGSLRVTF ASGMEIGLSS  1740
EPHILAGAVN PTLGKCNISL PGEHNANLIE WRQRKEQNKG NVSAFERRLR AHNRNLLSID  1800
FDHITRTGKI YDDHRKFTLR ILYDQTGRPI LWSPVSRYNE VNITYSPSGL VTFIQRGTWN  1860
EKMEYDQSGK IISRTWADGK IWSYTYLEKS VMLLLHSQRR YIFEYDQSDC LLSVTMPSMV  1920
RHSLQTMLSV GYYRNIYTPP DSSTSFIQDY SRDGRLLQTL HLGTGRRVLY KYTKQARLSE  1980
VLYDTTQVTL TYEESSGVIK TIHLMHDGFI CTIRYRQTGP LIGRQIFRFS EEGLVNARFD  2040
YSYNNFRVTS MQAVINETPL PIDLYRYVDV SGRTEQFGKF SVINYDLNQV ITTTVMKHTK  2100
IFSANGQVIE VQYEILKAIA YWMTIQYDNV GRHGNMCIRV GVDANITRYF YEYDADGQLQ  2160
TVSVNDKTQW RYSYDLNGDI NLLSHGKSAR LTPLRYDLRD RITRLGEIQY KMDEDGFLRQ  2220
RGNDIFEYNS NGLLQKAYNK ASGWTVQYYY DGLGRRVASK SSLGQHLQFF VDATANPIRV  2280
THLYNHTSSE ITSLYYDLQG HLIAMELSSG EEYYVACDNT GTPLAVFSSR GQVIKEILYT  2340
PYGDIYHDTY PDFQVIIGFH GGLYDFLTKL VHLGQRDYDV VAGRWTTAYH HIWKQLNLLP  2400
KPFNLYSFEN NYPVGKIQDV AKYTTDIRSW LELFGFQLHN VLPGFPKPEL ENLELTYELL  2460
RLQTKTQEWD PGKTILGIQC ELQKQLRNFI SLDQLPMTPR YNDGRCLEGG KQPRFAAVPS  2520
VFGKGIKFAI KDGIVTADII GVANEDSRRL AAILNNAHYL ENLHFTIEGR DTHYFIKLGS  2580
LEEDLVLIGN TGGRRILENG VNVTVSQMTS LLNGRTRRFA DIQLQHGALC FNIRYGTTVE  2640
EEKNHVLEIA RQRAVAQAWT KEQRRLQEGE EGIRAWTEGE KQQLLSTGRV QGYDGYFVLS  2700
VEQYLELSDS ANNIHFMRQS EIGRR
```

SEQ ID NO:231 PFD4 DNA SEQUENCE:
Nucleic Acid Accession #:   NM_000441

397

Coding sequence: 225-2567 (underlined sequences correspond to start and stop codons)

```
        1          11         21         31         41         51
        |          |          |          |          |          |
     CTCAGCCTTC CCGGTTCGGG AAAGGGGAAG AATGCAGGAG GGGTAGGATT TCTTTCCTGA    60
     TAGGATCGGT TGGGAAAGAC CGCAGCCTGT GTGTGTCTTT CCCTTCGACC AAGGTGTCTG   120
     TTGCTCCGTA AATAAAAACGT CCCACTGCCT TCTGAGAGCG CTATAAAGGC AGCGGAAGGG   180
     TAGTCCGCGG GGCATTCCGG GCGGGGCGCG AGCAGAGACA GGTCATGGCA GCGCCAGGCG    240
     GCAGGTCGGA GCCGCCGCAG CTCCCCGAGT ACAGCTGCAG CTACATGGTG TCGCGGCCGG    300
     TCTACAGCGA GCTCGCTTTC CAGCAACAGC ACGAGCGGCG CCTGCAGGAG CGCAAGACGC    360
     TGCGGGAGAG CCTGGCCAAG TGCTGCAGTT GTTCAAGAAA GAGAGCCTTT GGTGTGCTAA    420
     AGACTCTTGT GCCCATCTTG GAGTGGCTCC CCAAATACCG AGTCAAGGAA TGGCTGCTTA    480
     GTGACGTCAT TTCGGGAGTT AGTACTGGGC TAGTGGCCAC GCTGCAAGGG ATGGCATATG    540
     CCCTACTAGC TGCAGTTCCT GTCGGATATG GTCTCTACTC TGCTTTTTTC CCTATCCTGA    600
     CATACTTTAT CTTTGGAACA TCAAGACATA TCTCAGTTGG ACCTTTTCCA GTGGTGAGTT    660
     TAATGGTGGG ATCTGTTGTT CTGAGCATGG CCCCCGACGA ACACTTTCTC GTATCCAGCA    720
     GCAATGGAAC TGTATTAAAT ACTACTATGA TAGACACTGC AGCTAGAGAT ACAGCTAGAG    780
     TCCTGATTGC CAGTGCCCTG ACTCTGCTGG TTGGAATTAT ACAGTTGATA TTTGGTGGCT    840
     TGCAGATTGG ATTCATAGTG AGGTACTTGG CAGATCCTTT GGTTGGTGGC TTCACAACAG    900
     CTGCTGCCTT CCAAGTGCTG GTCTCACAGC TAAAGATTGT CCTCAATGTT TCAACCAAAA    960
     ACTACAATGG AGTTCTCTCT ATTATCTATA CGCTGGTTGA GATTTTTCAA AATATTGGTG   1020
     ATACCAATCT TGCTGATTTC ACTGCTGGAT TGCTCACCAT TGTCGTCTGT ATGGCAGTTA   1080
     AGGAATTAAA TGATCGGTTT AGACACAAAA TCCCAGTCCC TATTCCTATA GAAGTAATTG   1140
     TGACGATAAT TGCTACTGCC ATTTCATATG GAGCCAACCT GGAAAAAAAT TACAATGCTG   1200
     GCATTGTTAA ATCCATCCCA AGGGGGTTTT TGCCTCCTGA ACTTCCACCT GTGAGCTTGT   1260
     TCTCGGAGAT GCTGGCTGCA TCATTTTCCA TCGCTGTGGT GGCTTATGCT ATTGCAGTGT   1320
     CAGTAGGAAA AGTATATGCC ACCAAGTATG ATTACACCAT CGATGGGAAC CAGGAATTCA   1380
     TTGCCTTTGG GATCAGCAAC ATCTTCTCAG GATTCTTCTC TTGTTTTGTG GCCACCACTG   1440
     CTCTTTCCCG CACGGCCGTC CAGGAGAGCA CTGGAGGAAA GACACAGGTT GCTGGCATCA   1500
     TCTCTGCTGC GATTGTGATG ATCGCCATTC TTGCCCTGGG GAAGCTTCTG GAACCCTTGC   1560
     AGAAGTCGGT CTTGGCAGCT GTTGTAATTG CCAACCTGAA AGGGATGTTT ATGCAGCTGT   1620
     GTGACATTCC TCGTCTGTGG AGACAGAATA AGATTGATGC TGTTATCTGG GTGTTTACGT   1680
     GTATAGTGTC CATCATTCTG GGGCTGGATC TCGGTTTACT AGCTGGCCTT ATATTTGGAC   1740
     TGTTGACTGT GGTCCTGAGA GTTCAGTTTC CTTCTTGGAA TGGCCTTGGA AGCATCCCTA   1800
     GCACAGATAT CTACAAAAGT ACCAAGAATT ACAAAAACAT TGAAGAACCT CAAGGAGTGA   1860
     AGATTCTTAG ATTTTCCAGT CCTATTTTCT ATGGCAATGT CGATGGTTTT AAAAAATGTA   1920
     TCAAGTCCAC AGTTGGATTT GATGCCATTA GAGTATATAA TAAGAGGCTG AAAGCGCTGA   1980
     GGAAAATACA GAAACTAATA AAAAGTGGAC AATTAAGAGC AACAAAGAAT GGCATCATAA   2040
     GTGATGCTGT TTCAACAAAT AATGCTTTTG AGCCTGATGA GGATATTGAA GATCTGGAGG   2100
     AACTTGATAT CCCAACCAAG GAAATAGAGA TTCAAGTGGA TTGGAACTCT GAGCTTCCAG   2160
     TCAAAGTGAA CGTTCCCAAA GTGCCAATCC ATAGCCTTGT GCTTGACTGT GGAGCTATAT   2220
     CTTTCCTGGA CGTTGTTGGA GTGAGATCAC TGCGGGTGAT TGTCAAAGAA TTCCAAAGAA   2280
     TTGATGTGAA TGTGTATTTT GCATCACTTC AAGATTATGT GATAGAAAAG CTGGAGCAAT   2340
     GCGGGTTCTT TGACGACAAC ATTAGAAAGG ACACATTCTT TTTGACGGTC CATGATGCTA   2400
     TACTCTATCT ACAGAACCAA GTGAAATCTC AAGAGGGTCA AGGTTCCATT TTAGAAACGA   2460
     TCACTCTCAT TCAGGATTGT AAAGATACCC TTGAATTAAT AGAAACAGAG CTGACGGAAG   2520
     AAGAACTTGA TGTCCAGGAT GAGGCTATGC GTACACTTGC ATCCTGAAAG TGGGTTCGGG   2580
     AGGTCTCTAT GAGCAAGGAA TACAAGACAA AACTTCCTCA ATGCATTGAC TATTTCTTCA   2640
     GACTCAAAAC ACTCATTCTT TTTTCTATTA AGCCATTGAA AGAGAAGCAC TAAGACTGCT   2700
     TCTAGGCTTT ATTTATAAAA TAAACACCTT ATCCCTAACA TGGGCAAAAT GGCTAGAATT   2760
     ATTCAGACGA TTTGGCAGCG TCCAGGGTAA GCTGGTGTTA TAATACGCTG CTGATCTACA   2820
     TCACAGATTT GCTAATAATG TTCACGTGGG CCCTGGCATA TCTCTGTTCA GTTAGAGTGA   2880
     GTGCTGACCC AACAGCCTCT GTGGTCAAGC GAGTCACGAA TGATTAATCA TAAAGAAAAA   2940
     TCAGTTTTTG ACTGACCTGG ATATCCATGA GCTGCACTGA TCACCATGTA AGGTCACATT   3000
     TAGTAAATGC TGAAATAAAA TGATTAATGC ATTTATCAAT AAAAGCCTTT GAAAATACTT   3060
     TGGATAATAA ATTGGAGTTT TAAAAATGCA AATTTGCTTA GTATCTAATA ATGAAGTGTT   3120
     ATTACATATA GCCGGAATTG AGGATCTCTT TGATCCTGGA AATGGTTTAC CTAAAAGCTA   3180
     CAGAACCAGG CCAATATATT TTGAAATATT GATGCAGACA AATGAAAATAA TAAAGAGATT   3240
     TTCATGGTTT ATAAAAATCT TTTTTGATAT GATAATAATC ATGATCACAA CTGAGATCAA   3300
     AAAAATATAT GACAGATTAT TTTGTTTAAA AATGCAGTTT TAATTATCTT AGTCTATAGA   3360
     AATGATCATT GCATGGAGGC ATGTATAGGT ATGATCTGTG TAAAATCTGA CATAAAAACA   3420
     GTGCTATTCT GAGTGAAAAT TTTTTTGATG TGCTTACATA ACCATGGTGA TTAAAATGAG   3480
     TTTATATTTT TTCTCAAAAA TTTTTAGCAGT GTGTAAAGTA AGTAATCTTT AACTGAACTC   3540
     TGACCACTTA AAAAAAAATC TAAAAATTGA ACTACCTATA GTAGTCTGTG TTTAAAGTGA   3600
     ATTTTAAAG ACAAAGCATT CTAAATGAAC TCAATATAAA AACATTCATT TGGAATGTAC   3660
     ATACTGAAAA ATACAGGTTT TTTTGACCAA AAGTTTTTAT ATCTTTTCTT TTTATTTATT   3720
     TTTTTCCTAA GTGCCAACAA TTTTCTAGAT ATTATATACA ACACAGGCTT TGATCTTGGG   3780
     GACTTTTCCC ATATATTTCA CACTGGAGTG AATGAAGTTG TACTTCATTT CTAGAGAAAA   3840
     GTTATACCCA GGTCCCCAAT TGAGAATGTC TTGCTTGATT GAAAACGACA TCATCCCTTG   3900
     GTATACTCCA GGGATTGGTT TCAGGACCCC TGCATTTACC AAAATTTGTG CACACTCAAG   3960
     TCCTGCAGTC ACCCCTGCCT AAAGATAGAA TGGCTTCTCT GTTTTTCTTC TGAAATACAA   4020
     CCAGAAACAA TGTGTCTATT TCTGAAAGAA TAGGATTAAT GATCATACAA ATGGGTTAAT   4080
     CCTGAATTCT GGTTGTAAAT CTGGTTACAG CATAACTAGG ATTATAATGC TGCCTCATTT   4140
     TCACAGCACT ACTTGCTTAT ATTGACAACA AATCATCTCG CTAAAGAGTG AATGTAGGCC   4200
     AGGCGCGGTG GCTCATGCCT GTAATCCCAG CACTTTGGGA GGCCGAGGCG GGTGGATCAC   4260
     GAGGTCAGGA GATCGAGACC ATCCTGGCTA ACATGGTAAA ACCCCGTCTC TACTAAAAAT   4320
     AGAAAAAAAG AAATTAGCCT AGCGTGGTGG CTGGCGGGCG CCTGTAGTCC CAGCTATTTG   4380
     GGAGGCTAAG GCAGGAGAAT GGCGTGAACC CGGGAGGCGG AGCTTGCAGT GAGCCGAGGT   4440
     CGTGCCACTG CACTCCAGCC TGGGCGACAG AGCAAGACTC CGTCTCAAAA AAAAAAAAAA   4500
```

398

```
AAAAAAAAAA AGAGTGAATG TAATAGTCTT GCAGAAAATG AATGAATACC TTTGTTCAAT   4560
AAAGGAAATA TGCACTGCTC ACTTTTTTGA AGGAAATGCC AAAGTTACGT TTTACAACAA   4620
GGCTAGAGTT TGTAAATTCT GGGTTCATTT GTGATGACAT AAGTCAGCAA ACTGCGGGAA   4680
TACTGTCTCT TCTATGTATT TTGTGAATAG TAAGCATAAT TTTAGTTTTG TATTATCAAT   4740
GAAAATTTCA CTTGAAATTA AAGCTGCCTT TTGTTATATT TTTAACCTAT AGGATAAGAT   4800
TCCAGTATTG TATATGAGTT TTAACAAATT AAAAAAATCAA ATCATGTACA TTTGAAAATA   4860
TTTGCACACA TTTAAAAATA AATGTAAAGT TGTCTTTTAA ACTACTCGGA TGTGTCCTTT   4920
CTGAACAAAA
```

<u>SEQ ID NO:232 PFD4 Protein sequence:</u>
Protein Accession #:   O43511

```
1          11         21         31         41         51
|          |          |          |          |          |
MAAPGGRSEP PQLPEYSCSY MVSRPVYSEL AFQQQHERRL QERKTLRESL AKCCSCSRKR   60
AFGVLKTLVP ILEWLPKYRV KEWLLSDVIS GVSTGLVATL QGMAYALLAA VPVGYGLYSA  120
FFPILTYFIF GTSRHISVGP FPVVSLMVGS VVLSMAPDEH FLVSSSNGTV LNTTMIDTAA  180
RDTARVLIAS ALTLLVGIIQ LIFGGLQIGF IVRYLADPLV GGFTTAAAFQ VLVSQLKIVL  240
NVSTKNYNGV LSIIYTLVEI FQNIGDTNLA DFTAGLLTIV VCMAVKELND RFRHKIPVPI  300
PIEVIVTIIA TAISYGANLE KNYNAGIVKS IPRGFLPPEL PPVSLFSEML AASFSIAVVA  360
YAIAVSVGKV YATKYDYTID GNQEFIAFGI SNIFSGFFSC FVATTALSRT AVQESTGGKT  420
QVAGIISAAI VMIAILALGK LLEPLQKSVL AAVVIANLKG MFMQLCDIPR LWRQNKIDAV  480
IWVFTCIVSI ILGLDLGLLA GLIFGLLTVV LRVQFPSWNG LGSIPSTDIY KSTKNYKNIE  540
EPQGVKILRF SSPIFYGNVD GFKKCIKSTV GFDAIRVYNK RLKALRKIQK LIKSGQLRAT  600
KNGIISDAVS TNNAFEPDED IEDLEELDIP TKEIEIQVDW NSELPVKVNV PKVPIHSLVL  660
DCGAISFLDV VGVRSLRVIV KEFQRIDVNV YFASLQDYVI EKLEQCGFFD DNIRKDTFFL  720
TVHDAILYLQ NQVKSQEGQG SILETITLIQ DCKDTLELIE TELTEEELDV QDEAMRTLAS  780
QDEAMRTLAS
```

SEQ ID NO:233 PFH2 DNA SEQUENCE:

Nucleic Acid Accession #:   NM_016029
Coding sequence: 228-1097 (underlined sequences correspond to start and stop codons)

```
1          11         21         31         41         51
|          |          |          |          |          |
CTGCGATCCC GCAGGGCAGC GACGCGACTC TGGTGCGGGC CGTCTTCTTC CCCCCGAGCT   60
GGGCGTGCGC GGCCGCAATG AACTGGGAGC TGCTGCTGTG GCTGCTGGTG CTGTGCGCGC  120
TGCTCCTGCT CTTGGTGCAG CTGCTGCGCT TCCTGAGGGC TGACGGCGAC CTGACGCTAC  180
TATGGGCCGA GTGGCAGGGA CGACGCCCAG AATGGGAGCT GACTGATATG GTGGTGTGGG  240
TGACTGGAGC CTCGAGTGGA ATTGGTGAGG AGCTGGCTTA CCAGTTGTCT AAACTAGGAG  300
TTTCTCTTGT GCTGTCAGCC AGAAGAGTGC ATGAGCTGGA AAGGGTGAAA AGAAGATGCC  360
TAGAGAATGG CAATTTAAAA GAAAAAGATA TACTTGTTTT GCCCCTTGAC CTGACCGACA  420
CTGGTTCCCA TGAAGCGGCT ACCAAAGCTG TTCTCCAGGA GTTTGGTAGA ATCGACATTC  480
TGGTCAACAA TGGGTGGAATG TCCCAGCGTT CTCTGTGCAT GGATACCAGC TTGGATGTCT  540
ACAGAAAGCT AATAGAGCTT AACTACTTAG GGACGGTGTC CTTGACAAAA TGTGTTCTGC  600
CTCACATGAT CGAGAGGAAG CAAGGAAAGA TTGTTACTGT GAATAGCATC CTGGGTATCA  660
TATCTGTACC TCTTTCCATT GGATACTGTG CTAGCAAGCA TGCTCTCCGG GGTTTTTTTA  720
ATGGCCTTCG AACAGAACTT GCCACATACC CAGGTATAAT AGTTTCTAAC ATTTGCCCAG  780
GACCTGTGCA ATCAAATATT GTGGAGAATT CCCTAGCTGG AGAAGTCACA AAGACTATAG  840
GCAATAATGG AGACCAGTCC CACAAGATGA CAACCAGTCG TTGTGTGCGG CTGATGTTAA  900
TCAGCATGGC CAATGATTTG AAAGAAGTTT GGATCTCAGA ACAACCTTTC TTGTTAGTAA  960
CATATTTGTG GCAATACATG CCAACCTGGG CCTGGTGGAT AACCAACAAG ATGGGGAAGA 1020
AAAGGATTGA GAACTTTAAG AGTGGTGTGG ATGCAGACTC TTCTTATTTT AAAATCTTTA 1080
AGACAAAACA TGACTGAAAA GAGCACCTGT ACTTTTCAAG CCACTGGAGG GAGAAATGGA 1140
AAACATGAAA ACAGCAATCT TCTTATGCTT CTGAATAATC AAAGACTAAT TTGTGATTTT 1200
ACTTTTTAAT AGATATGACT TTGCTTCCAA CATGGAATGA AATAAAAAAT AAATAATAAA 1260
AGATTGCCAT GAATCTTGCA AA
```

<u>SEQ ID NO:234 PFH2 Protein sequence:</u>
Protein Accession #:   NP_057113

```
1          11         21         31         41         51
|          |          |          |          |          |
MNWELLLWLL VLCALLLLLV QLLRFLRADG DLTLLWAEWQ GRRPEWELTD MVVWVTGASS   60
GIGEELAYQL SKLGVSLVLS ARRVHELERV KRRCLENGNL KEKDILVLPL DLTDTGSHEA  120
ATKAVLQEFG RIDILVNNGG MSQRSLCMDT SLDVYRKLIE LNYLGTVSLT KCVLPHMIER  180
KQGKIVTVNS ILGIISVPLS IGYCASKHAL RGFFNGLRTE LATYPGIIVS NICPGPVQSN  240
IVENSLAGEV TKTIGNNGDQ SHKMTTSRCV RLMLISMAND LKEVVWISEQP FLLVTYLWQY  300
MPTWAWWITN KMGKKRIENF KSGVDADSSY FKIFKTKHD
```

SEQ ID NO:235 ACC5 DNA SEQUENCE

Nucleic Acid Accession #:   NM_000450

Coding sequence: 1-1833 (underlined sequences correspond to start and stop codons)

```
     1           11          21          31          41          51
     |           |           |           |           |           |
     ATGATTGCTT  CACAGTTTCT  CTCAGCTCTC  ACTTTGGTGC  TTCTCATTAA  AGAGAGTGGA     60
     GCCTGGTCTT  ACAACACCTC  CACGGAAGCT  ATGACTTATG  ATGAGGCCAG  TGCTTATTGT    120
     CAGCAAAGGT  ACACACACCT  GGTTGCAATT  CAAAACAAAG  AAGAGATTGA  GTACCTAAAC    180
     TCCATATTGA  GCTATTCACC  AAGTTATTAC  TGGATTGGAA  TCAGAAAAGT  CAACAATGTG    240
     TGGGTCTGGG  TAGGAACCCA  GAAACCTCTG  ACAGAAGAAG  CCAAGAACTG  GGCTCCAGGT    300
     GAACCCAACA  ATAGGCAAAA  AGATGAGGAC  TGCGTGGAGA  TCTACATCAA  GAGAGAAAAA    360
     GATGTGGGCA  TGTGGAATGA  TGAGAGGTGC  AGCAAGAAGA  AGCTTGCCCT  ATGCTACACA    420
     GCTGCCTGTA  CCAATACATC  CTGCAGTGGC  CACGGTGAAT  GTGTAGAGAC  CATCAATAAT    480
     TACACTTGCA  AGTGTGACCC  TGGCTTCAGT  GGACTCAAGT  GTGAGCAAAT  TGTGAACTGT    540
     ACAGCCCTGG  AATCCCCTGA  GCATGGAAGC  CTGGTTTGCA  GTCACCCACT  GGGAAACTTC    600
     AGCTACAATT  CTTCCTGCTC  TATCAGCTGT  GATAGGGGTT  ACCTGCCAAG  CAGCATGGAG    660
     ACCATGCAGT  GTATGTCCTC  TGGAGAATGG  AGTGCTCCTA  TTCCAGCCTG  CAATGTGGTT    720
     GAGTGTGATG  CTGTGACAAA  TCCAGCCAAT  GGGTTCGTGG  AATGTTTCCA  AAACCCTGGA    780
     AGCTTCCCAT  GGAACACAAC  CTGTACATTT  GACTGTGAAG  AAGGATTTGA  ACTAATGGGA    840
     GCCCAGAGCC  TTCAGTGTAC  CTCATCTGGG  AATTGGGACA  ACGAGAAGCC  AACGTGTAAA    900
     GCTGTGACAT  GCAGGGCCGT  CCGCCAGCCT  CAGAATGGCT  CTGTGAGGTG  CAGCCATTCC    960
     CCTGCTGGAG  AGTTCACCTT  CAAATCATCC  TGCAACTTCA  CCTGTGAGGA  AGGCTTCATG   1020
     TTGCAGGGAC  CAGCCCAGGT  TGAATGCACC  ACTCAAGGGC  AGTGGACACA  GCAAATCCCA   1080
     GTTTGTGAAG  CTTTCCAGTG  CACAGCCTTG  TCCAACCCCG  AGCGAGGCTA  CATGAATTGT   1140
     CTTCCTAGTG  CTTCTGGCAG  TTTCCGTTAT  GGGTCCAGCT  GTGAGTTCTC  CTGTGAGCAG   1200
     GGTTTTGTGT  TGAAGGGATC  CAAAAGGCTC  CAATGTGGCC  CCACAGGGGA  GTGGGACAAC   1260
     GAGAAGCCCA  CATGTGAAGC  TGTGAGATGC  GATGCTGTCC  ACCAGCCCCC  GAAGGGTTTG   1320
     GTGAGGTGTG  CTCATTCCCC  TATTGGAAAA  TTCACCTACA  AGTCCTCTTG  TGCCTTCAGC   1380
     TGTGAGGAGG  GATTTGAATT  ATATGGATCA  ACTCAACTTG  AGTGCACATC  TCAGGGACAA   1440
     TGGACAGAAG  AGGTTCCTTC  CTGCCAAGTG  GTAAAATGTT  CAAGCCTGGC  AGTTCCGGGA   1500
     AAGATCAACA  TGAGCTGCAG  TGGGGAGCCC  GTGTTTGGCA  CTGTGTGCAA  GTTCGCCTGT   1560
     CCTGAAGGAT  GGACGCTCAA  TGGCTCTGCA  GCTCGGACAT  GTGGAGCCAC  AGGACACTGG   1620
     TCTGGCCTGC  TACCTACCTG  TGAAGCTCCC  ACTGAGTCCA  ACATTCCCTT  GGTAGCTGGA   1680
     CTTTCTGCTG  CTGGACTCTC  CCTCCTGACA  TTAGCACCAT  TTCTCCTCTG  GCTTCGGAAA   1740
     TGCTTACGGA  AAGCAAAGAA  ATTTGTTCCT  GCCAGCAGCT  GCCAAAGCCT  TGAATCAGAC   1800
     GGAAGCTACC  AAAAGCCTTC  TTACATCCTT  TAA                                  1833
```

SEQ ID NO:236 ACC5 Protein sequence:
Protein Accession #:   NP_000441

```
     1           11          21          31          41          51
     |           |           |           |           |           |
     MIASQFLSAL  TLVLLIKESG  AWSYNTSTEA  MTYDEASAYC  QQRYTHLVAI  QNKEEIEYLN     60
     SILSYSPSYY  WIGIRKVNNV  WVWVGTQKPL  TEEAKNWAPG  EPNNRQKDED  CVEIYIKREK    120
     DVGMWNDERC  SKKKLALCYT  AACTNTSCSG  HGECVETINN  YTCKCDPGFS  GLKCEQIVNC    180
     TALESPEHGS  LVCSHPLGNF  SYNSSCSISC  DRGYLPSSME  TMQCMSSGEW  SAPIPACNVV    240
     ECDAVTNPAN  GFVECFQNPG  SFPWNTTCTF  DCEEGFELMG  AQSLQCTSSG  NWDNEKPTCK    300
     AVTCRAVRQP  QNGSVRCSHS  PAGEFTFKSS  CNFTCEEGFM  LQGPAQVECT  TQGQWTQQIP    360
     VCEAFQCTAL  SNPERGYMNC  LPSASGSFRY  GSSCEFSCEQ  GFVLKGSKRL  QCGPTGEWDN    420
     EKPTCEAVRC  DAVHQPPKGL  VRCAHSPIGE  FTYKSSCAFS  CEEGFELYGS  TQLECTSQGQ    480
     WTEEVPSCQV  VKCSSLAVPG  KINMSCSGEP  VFGTVCKFAC  PEGWTLNGSA  ARTCGATGHW    540
     SGLLPTCEAP  TESNIPLVAG  LSAAGLSLLT  LAPFLLWLRK  CLRKAKKFVP  ASSCQSLESD    600
     GSYQKPSYIL                                                                610
```

SEQ ID NO:237 PM28 DNA SEQUENCE

Nucleic Acid Accession #:   N51002
Coding sequence: 1-3793 (underlined sequences correspond to start and stop codons)

```
     1           11          21          31          41          51
     |           |           |           |           |           |
     ATGATGTGTG  AAGTGATGCC  CACGATTAAT  GAGGACACCC  CAATGAGCCA  AAGGGGGTCC     60
     CAAAGCAGTG  GCTCGGACTC  AGACTCCCAT  TTTGAGCAGC  TGATGGTGAA  TATGCTAGAT    120
     GAAAGGGATC  GTCTTCTAGA  CACCCTTCGG  GAGACCCAGG  AAAGCCTCTC  ACTTGCCCAG    180
     CAAAGACTTC  AGGATGTCAT  CTATGACCGA  GACTCACTCC  AGAGACAGCT  CAATTCAGCC    240
     CTGCCACAGG  ATATCGAATC  CCTAACAGGA  GGGCTGGCTG  GTTCTAAGGG  GGCTGATCCA    300
     CCGGAATTTG  CTGCACTGAC  AAAAGAATTA  AATGCCTGCA  GGGAACAACT  TCTAGAAAAG    360
     GAAGAAGAAA  TCTCTGAACT  TAAAGCTGAA  AGAAACAACA  CAAGACTATT  ACTGGAGCAT    420
     TTGGAGTGCC  TTGTGTCACG  ACATGAAAGA  TCACTAAGAA  TGACGGTGGT  AAAACGGCAA    480
     GCCCAGTCTC  CCTCAGGAGT  ATCCAGTGAA  GTTGAAGTTC  TCAAGGCACT  GAAATCTTTG    540
     TTTGAGCACC  ACAAGGCCTT  GGATGAAAAG  GTAAGGGGAG  CGACTGAGGGT  TTCTTTAGAA   600
     AGAGTCTCTG  CACTGGAAGA  AGAACTAGCT  GCTGCTAATC  AGGAGATTGT  TGCCTTGCGT    660
     GAACAAAATG  TTCATATACA  AAGAAAAAATG  GCATCAAGCG  AGGGATCCAC  AGAGTCAGAA   720
     CATCTTGAAG  GGATGGAACC  TGGACAGAAA  GTCCATGAGA  AGCGTTTGTC  CAATGGTTCT    780
     ATAGACTCAA  CCGATGAAAC  TAGTCAAATA  GTTGAACTAC  AAGAATTGCT  TGAAAAGCAA    840
     AACTATGAAA  TGGCCCAGAT  GAAAGAACGT  TTAGCAGCCC  TTTCTTCCCG  AGTGGGAGAG    900
     GTGGAACAGG  AAGCAGACAA  AGCAAGAAAG  GATCTCATTA  AAACAGAAGA  AATGAACACC    960
     AAGTATCAAA  GGGACATTAG  GGAGGCCATG  CGACAAAAGG  AAGATATGGA  AGAAAGAATT   1020
     ACAACCCTTG  AAAAGCGTTA  CCTCAGTGCT  CAGAGAGAAT  CTACCTCCAT  ACATGACATG   1080
```

```
AATGATAAAC TAGAAAATGA GTTAGCAAAT AAAGAAGCTA TCCTACGGCA GATGGAAGAG    1140
AAAAACAGAC AGTTACAAGA ACGTCTTGAG CTAGCTGAAC AAAAGTTGCA GCAGACCATG    1200
AGAAAGGCTG AAACCTTGCC TGAAGTAGAG GCTGAACTGG CTCAGAGAAT TGCAGCCCTA    1260
ACCAAGGCTG AAGAGAGACA TGGAAATATT GAAGAACGTA TGAGACATTT AGAGGGTCAA    1320
CTTGAAGAGA AGAATCAAGA ACTTCAAAGA GCTAGGCAAA GAGAGAAAAT GAATGAGGAG    1380
CATAACAAGA GATTATCGGA TACGGTTGAT AGACTTCTGA CTGAATCCAA TGAACGCCTA    1440
CAACTACACT TAAAGGAAAG AATGGCTGCT CTAGAAGAAA AGAATGTTTT AATTCAAGAA    1500
TCAGAAACTT TCAGAAAGAA TCTTGAAGAA TCTTTACATG ATAAGGAAAG ATTAGCAGAA    1560
GAAATTGAAA AGCTGAGATC TGAACTTGAC CAATTGAAAA TGAGAACTGG CTCTTTAATT    1620
GAACCCACAA TACCAAGAAC TCATCTAGAC ACCTCAGCTG AGTTGCGGTA CTCAGTGGGA    1680
TCCCTAGTGG ACAGCCAGTC TGATTACAGA ACAACTAAAG TAATAAGAAG ACCAAGGAGA    1740
GGCCGCATGG GTGTGCGAAG AGATGAGCCA AAGGTGAAAT CTCTTGGGGA TCACGAGTGG    1800
AATAGAACTC AACAGATTGG AGTACTAAGC AGCCACCCTT TTGAAAGTGA CACTGAAATG    1860
TCTGATATTG ATGATGATGA CAGAGAAACA ATTTTTAGCT CAATGGATCT TCTCTCTCCA    1920
AGTGGTCATT CCGATGCCCA GACGCTAGCC ATGATGCTTC AGGAACAATT GGATGCCATC    1980
AACAAAGAAA TCAGGCTAAT TCAGGAAGAA AAAGAATCTA CAGAGTTGCG TGCTGAAGAA    2040
ATTGAAAATA GAGTGGCTAG TGTGAGCCTC GAAGGCCTGA ATTTGGCAAG GGTCCACCCA    2100
GGTACCTCCA TTACTGCCTC TGTTACAGCT TCATCGCTGG CCAGTTCATC TCCCCCCAGT    2160
GGACACTCAA CTCCAAAGCT CACCCCTCGA AGCCCTGCCA GGGAAATGGA TCGGATGGGA    2220
GTCATGACAC TGCCAAGTGA TCTGAGGAAA CATCGGAGAA AGATTGCAGT TGTGGAAGAA    2280
GATGGTCGAG AGGACAAAGC AACAATTAAA TGTGAAACTT CTCCTCCTCC TACCCCTAGA    2340
GCCCTCAGAA TGACTCACAC TCTCCCTTCT TCCTACCACA ATGATGCTCG AAGTAGTTTA    2400
TCTGTCTCTC TTGAGCCAGA AAGCCTCGGG CTTGGTAGTG CCAACAGCAG CCAAGACTCT    2460
CTTCACAAAG CCCCCAAGAA GAAAGGAATC AAGTCTTCAA TAGGACGTTT GTTTGGTAAA    2520
AAAGAAAAAG CTCGACTTGG GCAGCTCCGA GGCTTTATGG AGACTGAAGC TGCAGCTCAG    2580
GAGTCCCTGG GGTTAGGCAA ACTCGGAACT CAAGCTGAGA AGGATCGAAG ACTAAAGAAA    2640
AAGCATGAAC TTCTTGAAGA AGCTCGGAGA AAGGGATTAC CTTTTGCCCA GTGGGATGGG    2700
CCAACTGTGG TCGCATGGCT AGAGCTTTGG TTGGGAATGC CTGCGTGGTA CGTGGCAGCC    2760
TGCCGAGCCA ACGTGAAGAG TGGTGCCATC ATGTCTGCTT TATCTGACAC TGAGATCCAG    2820
AGAGAAATTG GAATCAGCAA TCCACTGCAT CGCTTAAAAC TTCGATTAGC AATCCAGGAG    2880
ATGGTTTCCC TAACAAGTCC TTCAGCTCCT CCAACATCTC GAACTCCTTC AGGCAACGTT    2940
TGGGTGACTC ATGAAGAAAT GGAAAATCTT GCAGCTCCAG CAAAAACGAA AGAATCTGAG    3000
GAAGGAAGCT GGGCCCAGTG TCCGGTTTTT CTACAGACCC TGGCTTATGG AGATATGAAT    3060
CATGAGTGGA TTGGAAATGA ATGGCTTCCC AGCTTGGGGT TACCTCAGTA CAGAAGTTAC    3120
TTTATGGAAT GCTTGGTAGA TGCAAGAATG TTAGATCACC TAACAAAAAA AGATCTCCGT    3180
GTCCATTTAA AAATGGTGGA TAGTTTCCAT CGAACAAGTT TACAATATGG AATTATGTGC    3240
TTAAAGAGGT TGAATTATGA CAGAAAAGAA CTAGAAAGAA GACGGGAAGC AAGCCAACAT    3300
GAAATAAAAG ACGTGTTGGT GTGGAGCAAT GACCGAATTA TTCGCTGGAT ACAAGCAATT    3360
GGACTTCGAG AATATGCAAA TAATATACTT GAGAGCGGTG TGCATGGCTC ACTTATAGCC    3420
CTGGATGAAA ACTTTGACTA CAGCAGCTTA ACTTTATTAT TACAGATTCC AACACAGAAC    3480
ACCCAGGCAA GGCAGATTCT TGAAAGAGAA TACAATAACC TCTTGGCCCT GGGAACTGAA    3540
AGGCGACTGG ATGAAAGTGA TGACAAGAAC TTCAGACGTG GATCAACCTG GAGAAGGCAG    3600
TTTCCTCCTC GTGAAGTACA TGGAATCAGC ATGATGCCTG GGTCCTCAGA AACATTACCA    3660
GCTGGATTTA GGTTAACCAC AACCTCTGGG CAATCAAGAA AAATGACAAC AGATGTTGCT    3720
TCATCAAGAC TGCAGAGGTT AGACAACTCC ACTGTTCGCA CATACTCATG TCTCGAGTAA    3780
GCGGCCGCTT TAA
```

```
1          11         21         31         41         51
|          |          |          |          |          |
MMCEVMPTIN EDTPMSQRGS QSSGSDSDSH FEQLMVNMLD ERDRLLDTLR ETQESLSLAQ    60
QRLQDVIYDR DSLQRQLNSA LPQDIESLTG GLAGSKGADP PEFAALTKEL NACREQLLEK    120
EEEISELKAE RNNTRLLLEH LECLVSRHER SLRMTVVKRQ AQSPSGVSSE VEVLKALKSL    180
FEHHKALDEK VRERLRVSLE RVSALEEELA AANQEIVALR EQNVHIQRKM ASSEGSTESE    240
HLEGMEPGQK VHEKRLSNGS IDSTDETSQI VELQELLEKQ NYEMAQMKER LAALSSRVGE    300
VEQEAETARK DLIKTEEMNT KYQRDIREAM AQKEDMEERI TTLEKRYLSA QRESTSIHDM    360
NDKLENELAN KEAILRQMEE KNRQLQERLE LAEQKLQQTM RKAETLPEVE AELAQRIAAL    420
TKAEERHGNI EERMRHLEGQ LEEKNQELQR ARQREKMNEE HNKRLSDTVD RLLTESNERL    480
QLHLKERMAA LEEKNVLIQE SETFRKNLEE SLHDKERLAE EIEKLRSELD QLKMRTGSLI    540
EPTIPRTHLD TSAELRYSVG SLVDSQSDYR TTKVIRRPRR GRMGVRRDEP KVKSLGDHEW    600
NRTQQIGVLS SHPFESDTEM SDIDDDDRET IFSSMDLLSP SGHSDAQTLA MMLQEQLDAI    660
NKEIRLIQEE KESTELRAEE IENRVASVSL EGLNLARVHP GTSITASVTA SSLASSSPPS    720
GHSTPKLTPR SPAREMDRMG VMTLPSDLRK HRRKIAVVEE DGREDKATIK CETSPPPTPR    780
ALRMTHTLPS SYHNDARSSL SVSLEPESLG LGSANSSQDS LHKAPKKKGI KSSIGRLFGK    840
KEKARLGQLR GFMETEAAAQ ESLGLGKLGT QAEKDRRLKK KHELLEEARR KGLPFAQWDG    900
PTVVAWLELW LGMPAWYVAA CRANVKSGAI MSALSDTEIQ REIGISNPLH RLKLRLAIQE    960
MVSLTPSPAP PTSRTPSGNV WVTHEEMENL AAPAKTKESE EGSWAQCPVF LQTLAYGDMN    1020
HEWIGNEWLP SLGLPQYRSY FMECLVDARM LDHLTKKDLR VHLKMVDSFH RTSLQYGIMC    1080
LKRLNYDRKE LERRREASQH EIKDVLVWSN DRIIRWIQAI GLREYANNIL ESGVHGSLIA    1140
LDENFDYSSL TLLLQIPTQN TQARQILERE YNNLLALGTE RRLDESDDKN FRRGSTWRRQ    1200
FPPREVHGIS MMPGSSETLP AGFRLTTTSG QSRKMTTDVA SSRLQRLDNS TVRTYSCLE
```

SEQ ID NO:239 PCI4 DNA SEQUENCE

401

```
       1          11         21         31         41         51
       |          |          |          |          |          |
 5   ATGAGGCGAC TGAATCGGAA AAAAACTTTA AGTTTGGTAA AAGAGTTGGA TGCCTTTCCG    60
     AAGGTTCCTG AGAGCTATGT AGAGACTTCA GCCAGTGGAG GTACAGTTTC TCTAATAGCA   120
     TTTACAACTA TGGCTTTATT AACCATAATG GAATTCTCAG TATATCAAGA TACATGGATG   180
     AAGTATGAAT ACGAAGTAGA CAAGGATTTT TCTAGCAAAT TAAGAATTAA TATAGATATT   240
     ACTGTTGCCA TGAAGTGTCA ATATGTTGGA GCGGATGTAT TGGATTTAGC AGAAACAATG   300
10   GTTGCATCTG CAGATGGTTT AGTTTATGAA CCAACAGTAT TTGATCTTTC ACCACAGCAG   360
     AAAGAGTGGC AGAGGATGCT GCAGCTGATT CAGAGTAGGC TACAAGAAGA GCATTCACTT   420
     CAAGATGTGA TATTTAAAAG TGCTTTTAAA AGTACATCAA CAGCTCTTCC ACCAAGAGAA   480
     GATGATTCAT CACAGTCTCC AAATGCATGC AGAATTCATG GCCATCTATA TGTCAATAAA   540
     GTAGCAGGGA ATTTTCACAT AACAGTGGGC AAGGCAATTC CACATCCTCG TGGTCATGCA   600
15   CATTTGGCAG CACTTGTCAA CCATGAATCT TACAATTTTT CTCATAGAAT AGATCATTTG   660
     TCTTTTGGAG AGCTTGTTCC AGCAATTATT AATCCTTTAG ATGGAACTGA AAAAATTGCT   720
     ATAGATCACA ACCAGATGTT CCAATATTTT ATTACAGTTG TGCCAACAAA ACTACATACA   780
     TATAAAATAT CAGCAGACAC CCATCAGTTT TCTGTGACAG AAAGGGAACG TATCATTAAC   840
     CATGCTGCAG GCAGCCATGG AGTCTCTGGG ATATTTATGA AATATGATCT CAGTTCTCTT   900
20   ATGGTGACAG TTACTGAGGA GCACATGCCA TTCTGGCAGT TTTTTGTAAG ACTCTGTGGT   960
     ATTGTTGGAG GAATCTTTTC AACAACAGGC ATGTTACATG GAATTGGAAA ATTTATAGTT  1020
     GAAATAATTT GCTGTCGTTT CAGACTTGGA TCCTATAAAC CTGTCAATTC TGTTCCTTTT  1080
     GAGGATGGCC ACACAGACAA CCACTTACCT CTTTTAGAAA ATAATACACA TTGA
```

```
       1          11         21         31         41         51
       |          |          |          |          |          |
     MRRLNRKKTL SLVKELDAFP KVPESYVETS ASGGTVSLIA FTTMALLTIM EFSVYQDTWM    60
     KYEYEVDKDF SSKLRINIDI TVAMKCQYVG ADVLDLAETM VASADGLVYE PTVFDLSPQQ   120
     KEWQRMLQLI QSRLQEEHSL QDVIFKSAFK STSTALPPRE DDSSQSPNAC RIHGHLYVNK   180
     VAGNFHITVG KAIPHPRGHA HLAALVNHES YNFSHRIDHL SFGELVPAII NPLDGTEKIA   240
     IDHNQMFQYF ITVVPTKLHT YKISADTHQF SVTERERIIN HAAGSHGVSG IFMKYDLSSL   300
     MVTVTEEHMP FWQFFVRLCG IVGGIFSTTG MLHGIGKFIV EIICCRFRLG SYKPVNSVPF   360
     EDGHTDNHLP LLENNTH
```

```
     AATTCGCCCT TGCTTAATTA AGCATGTTTA CCTTCCTGTC ATCTGTCACT GCTGCTGTCA    60
     GTGGCCTCCT GGTGGGTTAT GAACTTGGGA TCATCTCTGG GGCTCTTCTT CAGATCAAAA   120
     CCTTATTAGC CCTGAGCTGC CATGAGCAGG AAATGGTTGT GAGCTCCCTC GTCATTGGAG   180
     CCCTCCTTGC CTCACTCACC GGAGGGGTCC TGATAGACAG ATATGGAAGA AGGACAGCAA   240
     TCATCTTGTC ATCCTGCCTG CTTGGACTCG GAAGCTTAGT CTTGATCCTC AGTTTATCCT   300
     ACACGGTTCT TATAGTGGGA CGCATTGCCA TAGGGGTTTC CATCTCCCTC TCTTCCATTG   360
     CCACTTGTGT TTACATCGCA GAGATTGCTC CTCAACACAG AAGAGGCCTT CTTGTGTCAC   420
     TGAATGAGCT GATGATTGTC ATCGGCATTC TTTCTGCCTA TATTTCAAAT TACGCATTTG   480
     CCAATGTTTT CCATGGCTGG AAGTACATGT TTGGTCTTGT GATTCCCTTG GGAGTTTGTAC   540
     AAGCAATTGC AATGTATTTT CTTCCTCCAA GCCCTCGGTT TCTGGTGATG AAAGGACAAG   600
     AGGGAGCTGC TAGCAAGGTT CTTGGAAGGT TAAGAGCACT CTCAGATACA ACTGAGGAAC   660
     TCACTGTGAT CAAATCCTCC CTGAAAGATG AATATCAGTA CAGTTTTTGG GATCTGTTTC   720
     GTTCAAAAGA CAACATGCGG ACCCGAATAA TGATAGGACT AACACTAGTA TTTTTTGTAC   780
     AAATCACTGG CCAACCAAAC ATATTGTTCT ATGCATCAAC TGTTTTGAAG TCAGTTGGAT   840
     TTCAAAGCAA TGAGGCAGCT AGCCTCGCCT CCACTGGGGT TGGAGTCGTC AAGGTCATTA   900
     GCACCATCCC TGCCACTCTT CTTGTAGACC ATGTCGGCAG CAAAACATTC CTCTGCATTG   960
     GCTCCTCTGT GATGGCAGCT TCGTTGGTGA CCATGGGCAT CGTAAATCTC AACATCCACA  1020
     TGAACTTCAC CCATATCTGC AGAAGCCACA ATTCTATCAA CCAGTCCTTG GATGAGTCTG  1080
     TGATTTATGG ACCAGGAAAC CTGTCAACCA ACAACAATAC TCTCAGAGAC CACTTCAAAG  1140
     GGATTTCTTC CCATAGCAGA AGCTCACTCA TGCCCCTGGG AAATGATGTG GATAAGAAGA  1200
     GGGAGACGAC CTCAGCATCC TTGCTAAATG CTGGATTAAG CCACACTGAA TACCAGATAG  1260
     TCACAGACCC TGGGGACGTC CCAGCTTTTT TGAAATGGCT GTCCTTAGCC AGCTTGCTTG  1320
     TTTATGTTGC TGCTTTTTCA ATTGGTCTAG GACCAATGCC CTGGCTGGTG CTCAGCGGA G  1380
     TCTTTCCTGG TGGGATCAGA GGACGAGCCA TGGCTTTAAC TTCTAGCATG AACTGGGGCA  1440
     TCAATCTCCT CATCTCGCTG ACATTTTTGA CTGTAACTGA TCTTATTGGC CTGCCATGGG  1500
     TGTGCTTTAT ATATACAATC ATGAGTCTAG ATCTTATTGG CCTGCCATGG GTGTGCATTG  1560
     TATATACAAT CATGAGTCTA GCATCCCTGC TTTTTGTTGT TATGTTTATA CCTGAGACAA  1620
     AGGGATGCTC TTTGGAACAA ATATCAATGG AGCTAGCAAA AGTGAACTAT GTGAAAAACA  1680
     ACATTTGTTT TATGAGTCAT CACCAAGAAG AATTAGTGCC AAAACAGCCT CAAAAAAGAA  1740
     AACCCCAGGA GCAGCTCTTG GAGTGTAACA AGCTGTGTGG TAGGGGCCAA TCCAGGCAGC  1800
     TTTCTCCAGA GACCTAATGG CCTCAACACC TTCTGAACGT GGATAGTGCC AGAACACTTA  1860
     GGAGGGTGTC TTTGGACCAA TGCATAGTTG CGACTCCTGT GCTCTCTTTT CAGTGTCATG  1920
     GAACTGGTTT TGAAAGAGAC ACTCTGAAATG ATAAAGACAG CCTTTAATCC CCCTCCTCMC  1980
     CAGAAGGAAC CTCAAAAGGT AGATGAGGTA CAAGGTCCTA AGTGATCTCT TTTTCTGAGC  2040
     AGGATATCAG GTTAAAAAAA AAAAGTTACT GGCTGGTTTA ATACTTTCTA CCTTCTTCAC  2100
     AGAGCAGCCT TTGAATAGAC TATGTCCTAG TGAAGACATC AACCTCCGCC TTAAGCTATG  2160
     TATGTATGGA GGCCAGTCGC AGCTTTATTA TGCAGACACA CAAGTGGTCT GGACATGAGG  2220
     GTACAGTTTC TGCCTACCAA GACACTACTT GCACTGGATC TTACGCAAAA AAGAACCAGA  2280
```

402

```
      ACACACAGTG TGGACAACTG CCCATATATT CTATCTAGAT TAGGAGAGGG TCCTGGCTAG 2340
      GATTTTAGTG GTAATTCCTA GTTACATTCA ACAAGTATAA AGATTATAGA GCTTATTTTA 2400
      TGAACTATAA ACTATAATTT AATGCAAAAT ATCCTTTTAT GAATTTCATG TTAATATTGT 2460
  5   GAAATATTAA AATAATTCCR CAATAGTTGA GAAAAATGAG CATTTTTTTC CATTTTTAAA 2520
      AAATGCATAG AAAAGACAAT TTTAAAATCC TGGGACCATA TTTATTTAGA AGTAGCTGTT 2580
      AGTAAAACAT TAGAAAAGGA GTCAGGCCAT TAGGTTATTT ATCCAAATCT CTAAGCAATT 2640
      AGGTTGAAGT TATTAAGTCA AGCCTAGAAA AGCTGCCTCC TTGTAAGGCT TTCATGACAA 2700
      TGTATAGTAA TCCACAGTGT CCAATTCTTC ACACTCCTCA GAAATATCAC TACCTCAGGT 2760
 10   TACGGTACAC AGGCTATAAT TGATGATGAT GTTCAGATAA CTGAAGACAC AATAAATGAC 2820
      ATTCAGACAT CAGGAMAAWW CCCTCATGTT CTTTTCTATG ATGGCCACCT GTACCAGCAA 2880
      CGTGGGTTTC ACCCACACAA CGATGAACTG TTCTCTTACT TCTCCAGTTG ATTTTAAAGA 2940
      CTTGTTAAGA GGTCTTACTA ATAAAATTTG GGTATGATAG AAAAWCCACA ATCAAAWCTT 3000
      GAACCAAATA ACATATTAAA TTACTAATAT TTAAGTGATG GAAGACACAC AAAAAACTTA 3060
 15   AAAGCACGAA CAACCTAACT TGAAAAAGAA TTTTAAAATA TGATTAACCT GAAGAAAAGA 3120
      GAATCCTAAG AGCCAAAGCT CCTTTTTATT TAGCTTGGAA TTTTCCTATT GGTTCCTAAC 3180
      AAACTGTCCC AATGTCATAT AAGGAAACAT GATCTATTAC ATTCCTTTAT AACAATGTGG 3240
      AGAGACTATA AACCTATGTA AGTAGTAAAA CTATATYAGA GACTCAGGAG ACTGACTAAA 3300
      AGGCCTGGAT CTGCAGTGTA TTATCTGTAT AAAAATTGGC AGGGGGAAGC TAAAAGGAAA 3360
 20   GGAGATTGGA GATCTCAATT CTATCATGGT GTATTTCATA CGCAAATCAG AGCATGCATT 3420
      GTTTTTTGTT TTTGGAAAGA GAAGGGAAGT GTGTTCTGCC CCATGTTTCC TTCCGTGTTT 3480
      ATAGTTCAAA CTCTATATAT ACTTCAGGTA TTTTTTGTTT AGCCCTTCAT TATAAATGGG 3540
      CAGGAAATTG TTTATCAACC TAGCCAGTTT ATTACTAGTG ACCTTGACTT CAGTATCTTG 3600
      AGCATTCTTT TATATTTTTC TTTTATTATC CTGAGTCTGT AACTAAACAA TTTTGTCTTC 3660
 25   AAATTTTTAT CCAATATCCA TTGCACCACA CCAAATCAAG CTTCTTGATT TTCAAAAATA 3720
      AAAAGGGGGA AATACTTACA ACTTGTACAT ATATATTCAC AGTTTTTATT TATAAAAAAA 3780
      ATTTACAGTA CTTATGGAGA GCCAGCAGAA GACATCAGAG CACTCACTTC TTCCCATCTT 3840
      TGTTAAGGTT AGCGAATTAC CCATGGCACAC TGTTAGGTGA GGCTCATTCG GCAGCCCTGA 3900
      AAACAAACCT GGTCACACTG TCTTTACCCT CTCCCTTCAG ATAAAGCACT TCGATTATCT 3960
 30   ATTGATCTGC CCAGTTTTCA AGTCATGCGA ATACTAAAAA GGTTACATCA TCTGGATCTG 4020
      TACCTTGGCT ATATAAGCAT GTTTTCCCCC TATTCTATGT TTCTTTTTTT GGTGAACATT 4080
      GAAAAACAGG AGGTGACTTA TTACTGTTAA TTAAAACTAA ATGAAAAATG TCAAGTCTTT 4140
      AAAACAGTGA GCTTGTAACT CTTTCATGTA ATTTTATTCT CTATGAATTT GGCTATCCTA 4200
      CTGAATCTTA AAATAAAGGA AATAAACACT TTTTTTTWAA AAAAAAGGAA AAATAMAARW 4260
 35   MWAAAAATCT CAATGAAATA TTTCACAAGA AGGAAAAA
```

Protein Accession #:   AAF91431

```
MFTFLSSVTA AVSGLLVGYE LGIISGALLQ IKTLLALSCH EQEMVVSSLV IGALLASLTG   60
GVLIDRYGRR TAIILSSCLL GLGSLVLILS LSYTVLIVGR IAIGVSISLS SIATCVYIAE  120
IAPQHRRGLL VSLNELMIVI GILSAYISNY AFANVFHGWK YMFGLVIPLG VLQAIAMYFL  180
PPSPRFLVMK GQEGAASKVL GRLRALSDTT EELTVIKSSL KDEYQYSFWD LFRSKDNMRT  240
RIMIGLTLVF FVQITGQPNI LFYASTVLKS VGFQSNEAAS LASTGVGVVK VISTIPATLL  300
VDHVGSKTFL CIGSSVMAAS LVTMGIVNLN IHMNFTHICR SHNSINQSLD ESVIYGPGNL  360
STNNNTLRDH FKGISSHSRS SLMPLRNDVD KRGETTSASL LNAGLSHTEY QIVTDPGDVP  420
AFLKWLSLAS LLVYVAAFSI GLGPMPWLVL SEIFPGGIRG RAMALTSSMN WGINLLISLT  480
FLTVTDLIGL PWVCFIYTIM SLDLIGLPWV CFIYTIMSLA SLLFVVMFIP ETKGCSLEQI  540
SMELAKVNYV KNNICFMSHH QEELVPKQPQ KRKPQEQLLE CNKLCGRGQS RQLSPET
```

Nucleic Acid Accession#:   AA172056

Coding sequence: 121-339 (underlined sequences correspond to start and stop codons)

```
TTTAGCCACC AGAGGANTTC TCTTGAAATA CCCAAAATCC ATCAGTATCT TGAATCATGC   60
TGGATTTTGA AGAATTCTTA AGAAGCCATG TAAAGGGGGC TCTCTGGCCT TGAAATAGTG  120
ATGTTTTTTA TACAGAAAGG AGAATGCAGA ATGGTCAGAC TATCATGCAC TGTTAAATTT  180
GATTTCAAGA AATTACAGGA AAACTTTCCA AAGTTCCATC TCACAGAANN TTATTTTNCC  240
AAGAATTCCA AGATAAGTTT AGTTTTATGG AAGACTTTTA TGTGGTTTTT ACTCACTCTT  300
CATCTCAGAC ATCGACAGAT GATTACATCA CTTATAGTTC TAGTAAATTT ATTAATATAA  360
AACTCAGAGA CATTCCAATA TCCACATTGC TTACACCATT AGGCATAGAT TCAGTGTCAG  420
CTATGACAAT TGAAAATGAG CTGTTTTGTG ATTTAAAGGT TTAAATTTCT CTAACCAAAC  480
TGCTTGATCC AGATGCAGGA CTGCAAATGT TAATATTTGT TCTGGAAGAA CAATCAAATA  540
AGACTTAAGA GGAAAGGGAA TGGCCACAAT CCACCTGAAA TTTTTTCTTA AAAAGTGTGC  600
AGCCTACTAA ATCAGAATGA AAATAGAAGT ACAAGATTAT AAACAAAATG CAATCAAACT  660
TTTCTTAAGC TTACCTAAAG TTATTTCATC TGAAAATTTC AAGCAACTTT GTTCAACATT  720
AAATTGACAA TCTAAACTAA CAAGTCTTTT GAATTTATGC ATGGTAGTAA ACATTCTCTC  780
TATTAACTTT ATTACCTAAG GCTAAACCTA AAATTTTTAA GCAAAATTAG AAAAATAGTC  840
TTCACTCATC AAAAAAATAAA GTTTGTTACA TTTAGTATTT TCCCAATAAA ATTGGTCGTT  900
CTTGGTTTTT TATTTGGAGA GTCTGTGCAA AATGTCACTA AAAATAAATT AGCACTAGAA  960
ATTATTTCTA AATACCAAA
```

Nucleic Acid Accession#:   X51405
Coding sequence: 3-1721 (underlined sequence corresponds to start and stop codon)

```
1          11          21          31          41          51
```

403

```
    |          |          |          |          |          |
AAATGGCGTG CCCGTCTCTC CGCCGGCCCC CTGCCTCGCA GTGGTTTCTC CTGCAGCTCC    60
CCTGGGCTCC GCGGCCAGTA GTGCAGCCCG TGGAGCCGCG GCTTTGCCCG TCTCCTCTGG   120
GTGGCCCCAG TGCGCGGGCT GACACTCATT CAGCCGGGGA AGGTGAGGCG AGTAGAGGCT   180
GGTGCGGAAC TTGCCGCCCC CAGCAGCGCC GGCGGGCTAA GCCCAGGGCC GGGCAGACAA   240
AAGAGGCCGC CCGCGTAGGA AGGCACGGCC GGCGGCGGCG GAGCGCAGCG ATGGCCGGGC   300
GAGGGGGCAG CGCGCTGCTG GCTCTGTGCG GGGCACTGGC TGCCTGCCGG TGGCTCCTGG   360
GCGCCGAAGC CCAGGAGCCC GGGGCGCCCG CGGCGGGCAT GAGGCGGCGC CGGCGGCTGC   420
AGCAAGAGGA CGGCATCTCC TTCGAGTACC ACCGCTACCC CGAGCTGCGC GAGGCGCTCG   480
TGTCCGTGTG GCTGCAGTGC ACCGCCATCA GCAGGATTTA CACGGTGGGG CGCAGCTTCG   540
AGGGCCGGGA GCTCCTGGTC ATCGAGCTGT CCGACAACCC TGGCGTCCAT GAGCCTGGTG   600
AGCCTGAATT TAAATACATT GGGAATATGC ATGGGAATGA GGCTGTTGGA CGAGAACTGC   660
TCATTTTCTT GGCCCAGTAC CTATGCAACG AATACCAGAA GGGGAACGAG ACAATTGTCA   720
ACCTGATCCA CAGTACCCGC ATTCACATCA TGCCTTCCCT GAACCCAGAT GGCTTTGAGA   780
AGGCAGCGTC TCAGCCTGGT GAACTCAAGG ACTGGTTTGT GGGTCGAAGC AATGCCCAGG   840
GAATAGATCT GAACCGGAAC TTTCCAGACC TGGATAGGAT AGTGTACGTG AATGAGAAAG   900
AAGGTGGTCC AAATAATCAT CTGTTGAAAA ATATGAAGAA AATTGTGGAT CAAAACACAA   960
AGCTTGCTCC TGAGACCAAG GCTGTCATTC ATTGGATTAT GGATATTCCT TTTGTGCTTT  1020
CTGCCAATCT CCATGGAGGA GACCTTGTGG CCAATTATCC ATATGATGAG ACGCGGAGTG  1080
GTAGTGCTCA CGAATACAGC TCCTCCCCAG ATGACGCCAT TTTCCAAAGC TTGGGCCCGGG  1140
CATACTCTTC TTTCAACCCG GCCATGTCTG ACCCCAATCG GCCACCATGT CGCAAGAATG  1200
ATGATGACAG CAGCTTTGTA GATGGAACCA CCAACGGTGG TGCTTGGTAC AGCGTACCTG  1260
GAGGGATGCA AGACTTCAAT TACCTTAGCA GCAACTGTTT TGAGATCACC GTGGAGCTTA  1320
GCTGTGAGAA GTTCCCACCT GAAGAGACTC TGAAGACCTA CTGGGAGGAT AACAAAAACT  1380
CCCTCATTAG CTACCTTGAG CAGATACACC GAGGAGTTAA AGGATTTGTC CGAGACCTTC  1440
AAGGTAACCC AATTGCGAAT GCCACCATCT CCGTGGAAGG AATAGACCAC GATGTTACAT  1500
CCGCAAAGGA TGGTGATTAC TGGAGATTGC TTATACCTGG AAACTATAAA CTTACAGCCT  1560
CAGCTCCAGG CTATCTGGCA ATAACAAAGA AAGTGGCAGT TCCTTACAGC CCTGCTGCTG  1620
GGGTTGATTT TGAACTGGAG TCATTTTCTG AAAGGAAAGA AGAGGAGAAG GAAGAATTGA  1680
TGGAATGGTG GAAAATGATG TCAGAAACTT TAAATTTTTA AAAAGGCTTC TAGTTAGCTG  1740
CTTTAAATCT ATCTATATAA TGTAGTATGA TGTAATGTGG TCTTTTTTTT AGATTTTGTG  1800
CAGTTAATAC TTAACATTGA TTTATTTTTT AATCATTTAA ATATTAATCA ACTTTCCTTA  1860
AAATAAATAG CCTCTTAGGT AAAAATATAA GAACTTGATA TATTTCATTC TCTTATATAG  1920
TATTCATTTT CCTACCTATA TTACACAAAA AAGTATAGAA AAGATTTAAG TAATTTTGCC  1980
ATCCTAGGCT TAAATGCAAT ATTCCTGGTA TTATTTACAA TGCAGAATTT TTTGAGTAAT  2040
TCTAGCTTTC AAAAATTAGT GAAGTTCTTT TACTGTAATT GGTGACAATG TCACATAATG  2100
AATGCTATTG AAAAGGTTAA CAGATACAGC TCGGAGTTGT GAGCACTCTA CTGCAAGACT  2160
TAAATAGTTC AGTATAAATT GTCGTTTTTT TCTTGTGCTG ACTAACTATA AGCATGATCT  2220
TGTTAATGCA TTTTTGATGG GAAGAAAAGG TACATGTTTA CAAAGAGGTT TTATGAAAAG  2280
AATAAAAATT GACTTCTTGC TTGTACATAT AGGAGCAATA CTATTATATT ATGTAGTCCG  2340
TTAACACTAC TTAAAAGTTT AGGGTTTTCT CTTGGTTGTA GAGTGGCCCA GAATTGCATT  2400
CTGAATGAAT AAAGGTTAAA AAAAAATCCC CAGTGAAAAA AAA
```

SEQ ID NO:245 PBQ8 Protein sequence

Protein Accession#:   P16870

```
MAGRGGSALL ALCGALAACG WLLGAEAQEP GAPAAGMRRR RRLQQEDGIS FEYHRYPELR   60
EALVSVWLQC TAISRIYTVG RSFEGRELLV IELSDNPGVH EPGEPEFKYI GNMHGNEAVG  120
RELLIFLAQY LCNEYQKGNE TIVNLIHSTR IHIMPSLNPD GFEKAASQPG ELKDWFVGRS  180
NAQGIDLNRN FPDLDRIVYV NEKEGGPNNH LLKNMKKIVD QNTKLAPETK AVIHWIMDIP  240
FVLSANLHGG DLVANYPYDE TRSGSAHEYS SSPDDAIFQS LARAYSSFNP AMSDPNRPPC  300
RKNDDDSSFV DGTTNGGAWY SVPGGMQDFN YLSSNCFEIT VELSCEKFPP EETLKTYWED  360
NKNSLISYLE QIHRGVKGFV RDLQGNPIAN·ATISVEGIDH DVTSAKDGDY WRLLIPGNYK  420
LTASAPGYLA ITKKVAVPYS PAAGVDFELE SFSERKEEEK EELMEWWKMM SETLNF
```

SEQ ID NO:246 PBY4 DNA sequence

Nucleic Acid Accession#:   AF038966

Coding sequence: 91-1107 (underlined sequence corresponds to start and stop codon)

```
1          11         21         31         41         51
|          |          |          |          |          |
GGGGCGACGT GAGCGCGCAG GGGGGCGGCG GCCTCGCCTC GTCTCTCTCT CTGCGCCTGG    60
GTCGGGTGGG TGACGCCGAG AGCCAGAGAG ATGTCGGATT TCGACAGTAA CCCGTTTGCC   120
GACCCGGATC TCAACAATCC CTTCAAGGAT CCATCAGTTA CACAAGTGAC AAGAAATGTT   180
CCACCAGGAC TTGATGAATA TAATCCATTC TCGGATTCTA GAACACCTCC ACCAGGCGGT   240
GTGAAGATGC CTAATGTACC CAATACACAA CCAGCAATAA TGAAACCAAC AGAGGAACAT   300
CCAGCTTATA CACAGATTGC AAAGGAACAT GCATTGGCCC AAGCTGAACT TCTTAAGCGC   360
CAAGAAGAAC TAGAAAGAAA AGCCGCAGAA TTAGATCGTC GGGAACGAGA AATGCAAAAC   420
CTCAGTCAAC ATGGTAGAAA AAATATTTGG CCACCTCTTC CTAGCAATTT TCCTGTCGGA   480
CCTTGTTTCT ATCAGGAATT TTCTGTAGAC ATTCCTGTAG AATTCCAAAA GACAGTAAAG   540
CTTATGTACT ACTTGTGGAT GTTCATGCA GTAACACTGT TTCTAAATAT CTTCGGATGC   600
TTGGCTTGGT TTGTGTTGA TTCTGCAAGA GCGGTTGATT TTGGATTGAG TATCCTGTGG   660
TTCTTGCTTT TTACTCCTTG TTCATTTGTC TGTTGGTACA GACCACTTTA TGGAGCTTTC   720
AGGAGTGACA GTTCATTTAG ATTCTTTGTA TTCTTCTTCG TCTATATTTG TCAGTTTGCT   780
GTACATGTAC TCCAAGCTGC AGGATTTCAT AACTGGGCA ATTGTGGTTG GATTTCATCC   840
CTTACTGGTC TCAACCAAAA TATTCCTGTT GGAATCATGA TGATAATCAT AGCAGCACTT   900
TTCACAGCAT CAGCAGTCAT CTCACTAGTT ATGTTCAAAA AAGTACATGG ACTATATCGC   960
ACAACAGGTG CTAGTTTTGA GAAGGCCCAA CAGGAGTTTG CAACAGGTGT GATGTCCAAC  1020
AAAACTGTCC AGACCGCAGC TGCAAATGCA GCTTCAACTG CAGCATCTAG TGCAGCTCAG  1080
```

```
AATGCTTTCA AGGGTAACCA GATTTAAGAA TCTTCAAACA ATACACTGTT ACCTTTTGAC  1140
TGTACCTTTT TCTCCAGTTA CTGTATTCTA CAAATATTTT TATGTTCAAA ACACACAGTA  1200
CAGACAGCAT GGATATTTCC TGTTCACTTG TGCATGGGCT AAAACCAGGA AAACTTCCTT  1260
GTCTTATTAC TTTACCTAAT AGTTTCTTAA TATTTCAGTG CCCCTTGCAG AAAAAATATT  1320
ACATGCTAAA TAAATATTCT CCATATTTTT GGGGGATGAC ATTCAGTGAA TTATTTCAGT  1380
GGTGACCCAC TGAAAATTAA TAATGGTACT TATGATTAAA AACGCATTTA ATACTAACTG  1440
CAGTAGTTCT TTCAAGAATC TTTAGAGATA AGGATTGCAC ATTGGAAAAG TAAACCATGT  1500
TTCATTCCTT TTTCCCTATT TATATTGAAA GAAATAGGCC AGCAGAGACT TAGGGATTTT  1560
AAATTGGCTT GCTTTTTAGC TGTTTCAGTC ACCAGTGAAG AGCCTATGTG CATTTTGTAG  1620
TAGATAATGT AAAATTTGTC ATCTTTTTCT TTTCTTTTTT TTAGAATAGC TGATATTTTG  1680
ATAACAATCT CTAATTTGCA TGGGCACCAC ATTTCTTATA TTAAAAGAAT TAGTGTTTTG  1740
GCTTCTGTAC TGCTTATGGT TGTAGGATTC AGGGGTTAAT GGAATCACAG AAATGATATT  1800
CTGCAAGAAT TTCTTTTAAA TAAAAAGTTT GGGGGTGCAA TATAAGAAGT TTATATAATA  1860
TGCAGTACAT TATCCAAAAG AGAAGGTAGT TAATGCAGTA GAAAGTAGTG GTAATAATTC  1920
CTTTTT
```

SEQ ID NO: 247 PBY4 Protein sequence:

Protein Accession #:

```
MSDFDSNPFA DPDLNNPFKD PSVTQVTRNV PPGLDEYNPF SDSRTPPPGG VKMPNVPNTQ   60
PAIMKPTEEH PAYTQIAKEH ALAQAELLKR QEELERKAAE LDRREREMQN LSQHGRKNIW  120
PPLPSNFPVG PCFYQEFSVD IPVEFQKTVK LMYYLWMFHA VTLFLNIFGC LAWFCVDSAR  180
AVDFGLSILW FLLFTPCSFV CWYRPLYGAF RSDSSFRFFV FFFVYICQFA VHVLQAAGFH  240
NWGNCGWISS LTGLNQNIPV GIMMIIIAAL FTASAVISLV MFKKVHGLYR TTGASFEKAQ  300
QEFATGVMSN KTVQTAAANA ASTAASSAAQ NAFKGNQI
```

SEQ ID NO:248 PBH2 DNA sequence

Nucleic Acid Accession#:   none found
Coding sequence: 1-613 (underlined sequence corresponds to start and stop codon)

```
ATGAGAGACA ATAAATCGTG TGCTTTTTTC ATGGGAAAGT TAAATGTTTG TTTTGAAGGC   60
ACAGTAATAG CAGGCTATTC AGTGTTTGCC ACTACCTGCA TCATTCATCT GGCTGTAGCT  120
AGTGCACTAC AATTTCCTAA AAAGTCTTCT CACCCTCACA GGACTGCTCT ACATCTGGCC  180
TCTGCCAATG GAAATTCAGA AGTAGTAAAA CTCCTGCGTC ACAGACGATG TCAACTTAAT  240
ATCCTTGACA ACAAAAAGAG GACAGCTCTG ACAAAGGCCG TACAATGCCA GGAAGATGAA  300
TGTGCGTTAA TGTTGCTGGA ACATGGCACT GATCCGAATA TTCCAGATGA GTATGGAAAT  360
ACCGCTCTAC ACTATGCTAT CTACAATGAA GATAAATTAA TGGCCAAAGC ACTGCTCTTA  420
TACGGTGCTG ATATCGAATC AAAAAACAAG CATGGCCTCA CACCACTGTT ACTTGGTGTA  480
CATGAGCAAA AACAGCAAGT GGTGAAATTT TTAATCAAGA AAAAAGCAAA TTTAAATGCA  540
CTGGATAGAT ATGGAAGGTG TGTGACCTTG GGAACGTTAT TTACCACCAA ATATGTTGTC  600
ATATATGAAA AGTAG
```

SEQ ID NO:249 PBH2 Protein sequence:

Protein Accession #:   none found

```
MRDNKSCAFF MGKLNVCFEG TVIAGYSVFA TTCIIHLAVA SALQFPKKSS HPHRTALHLA   60
SANGNSEVVK LLLDRRCQLN ILDNKKRTAL TKAVQCQEDE CALMLLEHGT DPNIPDEYGN  120
TALHYAIYNE DKLMAKALLL YGADIESKNK HGLTPLLLGV HEQKQQVVKF LIKKKANLNA  180
LDRYGRCVTL GTLFTTKYVV IYEK
```

SEQ ID NO:250 PBJ1 DNA sequence

Nucleic Acid Accession#:   XM_005829
Coding sequence: 1-3043 (underlined sequence corresponds to start and stop codon)

```
ATGGTGATCA TCTATCTTTC TTTCTGCAAT TATTACATGG AGTTCTACAG AGAAGAGCTT   60
CCCCACATTG ACTATTTGAT TGACATTCAG TTTGCAACAG GAAAGGTTAC TCAGCCGGGA  120
GAGGACACTT CCTACCATCA ATGCGCTCAG CTTGAAGCCA GAGACGAAGG CACCGACAGT  180
TTATTATTAA ACAATGGCAG CAGCGCCACG CTGAAGACAC GAACGCGCTG TTATGGAACC  240
CCCAGAGGTC TCCCCCATCG TAGCCTGCTC CAGCCGACTC CGCCCACATG TAAAACGAAG  300
ATCAGGAGCA GATTTGAAGA ATTACAAAGT GAATTGGTGC CAGTCAGCAT GTCAGAGACA  360
GACCACATAG CCTCTACTTC CTCTGATAAA AATGTTGGGA AAACACCTGA ATTAAAGGAA  420
GACTCATGCA ACTTGTTTTC TGGCAATGAA AGCAGCAAAT TAGAAAATGA GTCCAAACTA  480
TTGTCATTAA ACACTGATAA AACTTTATGT CAACCTAATG AGCATAATAA TCGAATTGAA  540
GCCCAGGAAA ATTATATTCC AGATCATGGT GGAGGTGAGG ATTCTTGTGC CAAAACAGAC  600
ACAGGCTCAG AAAATTCTGA ACAAATAGCT AATTTTCCTA GTGGAAATTT TGCTAAACAT  660
ATTTCAAAAA CAAATGAAAC AGAACAGAAA GTAACACAAA TATTGGTGGA ATTAAGGTCA  720
TCTACATTTC CAGAATCAGC TAATGAAAAG ACTTATTCAG AAAGCCCCTA TGATACAGAC  780
TGCACCAAGA AATTTATTTC AAAAATAAAG AGCGTTTCAG CATCAGAGGA TTTGTTGGAA  840
GAAATAGAAT CTGAGCTCTT ATCTACGGAG TTTGCAGAAC ATCGAGTACC AAATGGAATG  900
AATAAGGGAG AACATGCATT AGTTCTGTTT GAAAAGTGTG TGCAAGATAA ATATTTGCAG  960
CAGGAACATA TCATAAAAAA GTTAATTAAA GAAAATAAGA AGCATCAGGA GCTCTTCGTA 1020
GACATTTGTT CAGAAAAAGA CAATTTAAGA GAAGAACTAA AGAAAGAAC AGAAACTGAG 1080
AAGCAGCATA TGAACACAAT TAAACAGTTA GAATCAAGAA TAGAAGAACT TAATAAAGAA 1140
GTTAAAGCTT CCAGAGATCA ACTAATAGCT CAAGACGTTA CAGCTAAAAA TGCAGTTCAG 1200
```

405

CAGTTACACA AAGAGATGGC CCAACGGATG GAACAGGCCA ACAAGAAATG TGAAGAGGCA 1260
CGCCAAGAAA AAGAAGCAAT GGTAATGAAA TATGTAAGAG GTGAGAAGGA ATCTTTAGAT 1320
CTTCGAAAGG AAAAAGAGAC ACTTGAGAAA AAACTTAGAG ATGCAAATAA GGAACTTGAG 1380
AAAAACACTA ACAAAATTAA GCAGCTTTCT CAGGAGAGAG GACGGTTGCA CCAGCTGTAT 1440
GAAACTAAGG AAGGCGAAAC GACTAGACTC ATCAGAGAAA TAGACAAATT AAAGGAAGAC 1500
ATTAACTCTC ACGTCATCAA AGTAAAGTGG GCACAAAACA AATTAAAAGC TGAAATGGAT 1560
TCACACAAGG AAACCAAAGA TAAACTCAAA GAAACAACAA CAAAATTAAC ACAAGCAAAG 1620
GAAGAAGCAG ATCAGATACG AAAAAACTGT CAGGATATGA TAAAAACATA TCAGGAGTCA 1680
GAAGAAATTA AATCAAATGA GCTTGATGCA AAGCTTAGAG TCACAAAAGG AGAACTTGAA 1740
AAACAAATGC AAGAAAAATC TGACCAGCTA GAGATGCATC ATGCCAAAAT AAAGGAACTA 1800
GAAGATCTGA AGAGAACATT TAAGGAGGGT ATGGATGAGT TAAGAACACT GAGAACAAAG 1860
GTGAAATGTC TAGAAGATGA ACGATTAAGA ACAGAAGATG AATTATCAAA ATATAAGGAA 1920
ATTATTAATC GCCAAAAAGC TGAAATTCAG AATTTATTGG ACAAGGTGAA AACTGCAGAT 1980
CAGCTACAGG AGCAGCTTCA AAGAGGTAAG CAAGAAATTG AAAATTTGAA AGAAGAAGTG 2040
GAAAGTCTTA ATTCTTTGAT TAATGACCTA CAAAAAAGCA TCGAAGGACG TAGGAAAAGA 2100
GAATCTGAGC TGCTGCTGTT TACAGAAAGG CTCACTAGTA AGAATGCACA GCTTCAGTCT 2160
GAATCCAATT CTTTGCAGTC ACAATTTGAT AAAGTTTCCT GTAGTGAAAG TCAGTTACAA 2220
AGCCAGTGTG AACAAATGAA ACAGACAAAT ATTAATTTGG AAAGTAGGTT GTTGAAAGAC 2280
GAAGAACTGC GAAAAGAGGA AGTCCAAACT CTGCAAGCTG AACTCGCTTG TAGACAAACA 2340
GAAGTTAAAG CATTGAGTAC CCAGGTAGAA GAATTAAAAG ATGAGTTAGT AACTCAGAGA 2400
CGTAAACATG CCTCTAGTAT CAAGGATCTC ACCAAACAAC TTCAGCAAGC ACGAAGAAAA 2460
TTAGATCAGG TTGAGAGTGG AAGCTATGAC AAAGAAGTCA GCAGCATGGG AAGTCGTTCT 2520
AGTTCATCAG GGTCCCTGAA TGCTCGAAGC AGTGCAGAAG ATCGATCTCC AGAAAATACT 2580
GGGTCCTCAG TAGCTGTGGA TAACTTTCCA CAAGTAGATA AGGCCATGTT GATTGAGAGA 2640
ATAGTTAGGC TGCAAAAAGC ACATGCCCGG AAAAATGAAA AGATAGAATT TATGGAGGAC 2700
CACATCAAAC AACTGGTGGA AGAAATTAGG AAAAAAACAA AAATAATTCA AAGTTATATT 2760
TTACGAGAAG AATCAGGCAC ACTTTCTTCA GAGGCATCTG ATTTTAACAA AGTTCATTTA 2820
AGTAGACGGG GTGGCATCAT GGCATCTTTA TATACATCCC ATCCAGCTGA CAATGGATTA 2880
ACATTGGAGC TCTCTTTGGA AATCAACCGA AAATTACAGG CTGTTTTGGA GGATACGTTA 2940
CTAAAAAATA TTACTTTGAA GGAAAATCTA CAAACACTTG AACAGAAAT AGAACGTCTT 3000
ATTAAACACC AGCATGAACT AGAACAGAGG ACAAAGAAAA CCTAAAACAA GCCTCTTGCT 3060
CAGTAAAGAG ACAAAAGCCA CACAGGAGTA GGTGCCACTG ACCTCTATTG TTGGAGACTT 3120
TGTTCCACTT TTTGTTTCAG CCAGTAAAAA TATTGTTTTG CTTCATCTGT ACACAAAAAA 3180
ATACCCTTTT ACAAATATGAA TGCATTGCTG TATATACTGT AAGACTGAAA GCTTTGATGA 3240
AATTTGTTTT TGTATGGTGC AATATGACAG CCTGTCATTG AATCTAAACA ACTTAATTTG 3300
CTTGTATTCA TAAGAAGTGT TGAACATTAC AAGGGCTTTT AT

MVIIYLSFCN YYMEFYREEL PHIDYLIDIQ FATGKVTQPG EDTSYHQCAQ LEARDEGTDS 60
LLLNNGSSAT LKTRTRCYGT PRGLPHRSLL QPTPPTCKTK IRSRFEELQS ELVPVSMSET 120
DHIASTSSDK NVGKTPELKE DSCNLFSGNE SSKLENESKL LSLNTDKTLC QPNEHNNRIE 180
AQENYIPDHG GGEDSCAKTD TGSENSEQIA NFPSGNFAKH ISKTNETEQK VTQILVELRS 240
STFPESANEK TYSESPYDTD CTKKFISKIK SVSASEDLLE EIESELLSTE FAEHRVPNGM 300
NKGEHALVLF EKCVQDKYLQ QEHIIKKLIK ENKKHQELFV DICSEKDNLR EELKKRTETE 360
KQHMNTIKQL ESRIEELNKE VKASRDQLIA QDVTAKNAVQ QLHKEMAQRM EQANKKCEEA 420
RQEKEAMVMK YVRGEKESLD LRKEKETLEK KLRDANKELE KNTNKIKQLS QEKGRLHQLY 480
ETKEGETTRL IREIDKLKED INSHVIKVKW AQNKLKAEMD SHKETKDKLK ETTTKLTQAK 540
EEADQIRKNC QDMIKTYQES EEIKSNELDA KLRVTKGELE KQMQEKSDQL EMHHAKIKEL 600
EDLKRTFKEG MDELRTLRTK VKCLEDERLR TEDELSKYKE IINRQKAEIQ NLLDKVKTAD 660
QLQEQLQRGK QEIENLKEEV ESLNSLINDL QKDIEGSRKR ESELLLFTER LTSKNAQLQS 720
ESNSLQSQFD KVSCSESQLQ SQCEQMKQTN INLESRLLKE EELRKEEVQT LQAELACRQT 780
EVKALSTQVE ELKDELVTQR RKHASSIKDL TKQLQQARRK LDQVESGSYD KEVSSMGSRS 840
SSSGSLNARS SAEDRSPENT GSSVAVDNFP QVDKAMLIER IVRLQKAHAR KNEKIEFMED 900
HIKQLVEEIR KKTKIIQSYI LREESGTLSS EASDFNKVHL SRRGGIMASL YTSHPADNGL 960
TLELSLEINR KLQAVLEDTL LKNITLKENL QTLGTEIERL IKHQHELEQR TKKT

```
1          11         21         31         41         51
|          |          |          |          |          |
TTGCCGTGAA GGGCTGTGCG GTTCCCGTGC GCGCCGGAGC CTGCTGTGGC CTCTTATGCA   60
CTCCACCACC CCCATCAGCT CCCTCTTCTC CTTCACCAGC CCCGCAGTGA AGAGACTGCT  120
AGGCTGGAAG CAAGGAGATG AAGAGGAAAA GTGGGCAGAG AAGGCAGTGG ACTCTCTAGT  180
GAAGAAGTTA AAGAAGAAGA AGGGAGCCAT GGACGAGCTG GAGAGGGCTC TCAGCTGCCC  240
GGGGCAGCCC AGCAAATGCG TCACGATTCC CCGCTCCCTG GACGGGCGGC TGCAGGTGTC  300
CCACCGCAAG GGCCTGCCCC ATGTGATTTA CTGTCGCGTG TGGCGCTGGC CGGATCTGCA  360
GTCCCACCAC GAGCTGAAGC CGCTGGAGTG CTGTGAGTTC CCATTTGGCT CCAAGCAGAA  420
AGAAGTGTGC ATTAACCCTT ACCACTACCG CCGGGTGGAG ACTCCAGTAC TGCCTCCTGT  480
GCTCGTGCCA AGACACAGTG AATATAACCC CCAGCTCAGC CTCCTGGCCA AGTTCCGCAG  540
CGCCTCCCTG CACAGTGAGC CACTCATGCC ACACAACGCC ACCTATCCTG ACTCTTTCCA  600
GCAGCCTCCG TGCTCTGCAC TCCCTCCCTC ACCCAGCCAC GCGTTCTCCC AGTCCCCGTG  660
CACGGCCAGC TACCCTCACT CCCCAGGAAG TCCTTCTGAG CCAGAGAGTC CCTATCAACA  720
CTCAGTTGAC ACACCACCCC TGCCTTATCA TGCCACAGAA GCCTCTGAGA CCCAGAGTGG  780
```

```
CCAACCTGTA GATGCCACAG CTGATAGACA TGTAGTGCTA TCGATACCAA ATGGAGACTT    840
TCGACCAGTT TGTTACGAGG AGCCCCAGCA CTGGTGCTCG GTCGCCTACT ATGAACTGAA    900
CAACCGAGTT GGGGAGACAT TCCAGGCTTC CTCCCGAAGT GTGCTCATAG ATGGGTTCAC    960
CGACCCTTCA AATAACAGGA ACAGATTCTG TCTTGGACTT CTTTCTAATG TAAACAGAAA   1020
CTCAACGATA GAAAATACCA GGAGACATAT AGGAAAGGGT GTGCACTTGT ACTACGTCGG   1080
GGGAGAGGTG TATGCCGAGT GCGTGAGTGA CAGCAGCATC TTTGTGCAGA GCCGGAACTG   1140
CAACTATCAA CACGGCTTCC ACCCAGCTAC CGTCTGCAAG ATCCCCAGCG GCTGCAGCCT   1200
CAAGGTCTTC AACAACCAGC TCTTCGCTCA GCTCCTGGCC CAGTCAGTTC ACCACGGCTT   1260
TGAAGTCGTG TATGAACTGA CCAAGATGTG TACTATCCGG ATGAGTTTTG TTAAGGGTTG   1320
GGGTGCTGAG TATCATCGCC AGGATGTCAC CAGCACCCCC TGCTGGATTG AGATTCATCT   1380
TCATGGGCCA CTGCAGTGGC TGGACAAAGT TCTGACTCAG ATGGGCTCTC CACATAACCC   1440
CATTTCTTCA GTGTCT TAAC AGTCATGTCT TAAGCTGCAT TTCCATAGGA T
```

SEQ ID NO:253 PBJ6 Protein sequence:
Protein Accession #:   NP_005896

```
MHSTTPISSL FSFTSPAVKR LLGWKQGDEE EKWAEKAVDS LVKKLKKKKG AMDELERALS    60
CPGQPSKCVT IPRSLDGRLQ VSHRKGLPHV IYCRVWRWPD LQSHHELKPL ECCEFPFGSK   120
QKEVCINPYH YRRVETPVLP PVLVPRHSEY NPQLSLLAKF RSASLHSEPL MPHNATYPDS   180
FQQPPCSALP PSPSHAFSQS PCTASYPHSP GSPSEPESPY QHSVDTPPLP YHATEASETQ   240
SGQPVDATAD RHVVLSIPNG DFRPVCYEEP QHWCSVAYYE LNNRVGETFQ ASSRSVLIDG   300
FTDPSNNRNR FCLGLLSNVN RNSTIENTRR HIGKGVHLYY VGGEVYAECV SDSSIFVQSR   360
NCNYQHGFHP ATVCKIPSGC SLKVFNNQLF AQLLAQSVHH GFEVVYELTK MCTIRMSFVK   420
GWGAEYHRQD VTSTPCWIEI HLHGPLQWLD KVLTQMGSPH NPISSVS
```

SEQ ID NO:254 PBJ8 DNA sequence
Nucleic Acid Accession#:   AB04684
Coding sequence: 472-4377 (underlined sequence corresponds to start and stop codon)

```
  1          11         21         31         41         51
  |          |          |          |          |          |
TGCAGGTTTG CAGGGTCTGA GATTACTTGG GCTTTTCCTG CCTTTTTCTT TTGCTTAAGG     60
GATGGACAAG GAGCTGAGAT TTATGACCCT TATTAGAGAA AAAAATGTGC CTTGCTAGGG    120
TGGGGACACT TGGTTGATGC AGTCTCTCTC TCTCTTTCTC GGTGTTTATA ACAAAACAAA    180
ACCAAAATGA ACTGAGGGGT TTGTAATGGT AGTTTGTTTG TTGCTGGAGA ATGCTACTTT    240
GCATGCTTTT TTTCTCTTGC AGGGTATGTT CTGTCTTGTG CTTTTTCTTT TAGAAGCTAC    300
TAAAGGGTGT TGGGGATGCT TCTGACTATT ATGAAGGCCA AAAGGCCTGT TGACTGGGGC    360
TGCTTTTAAC CCTTTCCTAT TTGCTGAGAA TGCAGCCGTG TGACAGTAAC TGAACATTGG    420
TCTAAAGTCT TTCCAAAAGG TCAAGGTTCA CAAGAACATC TGCTCAAATT AATGACCATG    480
GGGGATATGA AGACCCCAGA CTTTGATGAC CTCCTGGCAG CATTTGACAT CCCAGATATG    540
GTCGATCCTA AAGCAGCTAT TGAGTCTGGA CACGATGACC ATGAAAGCCA CATGAAGCAG    600
AATGCTCACG GAGAGGATGA CTCCCACGCA CCATCATCTT CTGATGTGGG TGTCAGCGTT    660
ATCGTCAAGA ATGTTCGGAA CATTGACTCT TCCGAGGGCG GGGAGAAAGA CGGCCACAAC    720
CCCACTGGCA ATGGCTTACA TAATGGGTTT CTCACAGCAT CCTCCCTTGA CAGTTACAGT    780
AAAGATGGAG CAAAGTCCTT GAAAGGAGAT GTGCCTGCCT CTGAGGTGAC ACTGAAAGAC    840
TCGACATTCA GCCAGTTTAG CCCGATCTCC AGTGCTGAAG AGTTTGATGA CGACGAGAAG    900
ATTGAGGTGG ATGACCCCCC TGACAAGGAG GACATGCGAT CAAGCTTCAG GTCGAATGTG    960
TTGACGGGGT CGGCTCCCCA GCAGGACTAC GATAAGCTGA AGGCACTCGG AGGGGAAAAC   1020
TCCAGCAAAA CTGGACTCTC TACGTCAGGC AATGTGGAGA AAAACAAAGC TGTTAAGAGA   1080
GAAACAGAAG CCAGTTCTAT AAACCTGAGT GTTTATGAAC CTTTTAAAGT CAGAAAAGCA   1140
GAGGATAAAT TGAAGGAAAG CTCTGACAAG GTGCTGGAAA ACAGAGTCCT AGATGGGAAG   1200
CTGAGCTCCG AGAAGAATGA CACCAGCCTC CCCAGCGTTG CGCCATCAAA GACAAAGTCG   1260
TCCTCCAAGC TCTCGTCCTG CATCGCTGCC ATCGCGGCTC TCAGCGCTAA AAAGGCGGCT   1320
TCAGACTCCT GCAAAGAACC AGTGGCCAAT TCGAGGGAAT CCTCCCCGTT ACCAAAAGAA   1380
GTAAATGACA GTCCGAGAGC CGCTGACAAG TCTCCTGAAT CCCAGAATCT CATCGACGGG   1440
ACCAAAAAAC CATCCCTGAA GCAACCGGAT AGTCCCAGAA GCATCTCAAG TGAGAACAGC   1500
AGCAAAGGAT CCCCGTCCTC TCCCGCAGGG TCCACACCAG CAATCCCCAA AGTCCGCATA   1560
AAAACCATTA AGACATCTTC TGGGGAAATC AAGAGAACAG TGACCAGGGT ATTGCCAGAA   1620
GTGGATCTTG ACTCTGGAAA GAAACCTTCC GAGCAGACAG CGTCCGTGAT GGCCTCTGTG   1680
ACATCCCTTC TGTCGTCTCC AGCATCAGCC GCCGTCCTTT CCTCTCCCCC CAGGGCGCCT   1740
CTCCAGTCTG CGGTCGTGAC CAATGCAGTT TCCCCTGCAG AGCTCACCCC CAAACAGGTC   1800
ACAATCAAGC CTGTGGCTAC TGCTTTCCTC CCAGTGTCTG CTGTGAAGAC GGCAGGATCC   1860
CAAGTCATTA ATTTGAAGCT CGCTAACAAC ACCACGGTGA AAGCCACGGT CATATCTGCT   1920
GCCTCTGTCC AGAGTGCCAG CAGCGCCATC ATTAAAGCTG CCAACGCCAT CCAGCAGCAA   1980
ACTGTCGTGG TGCCGGCATC CAGCCTGGCC AATGCCAAAC TCGTGCCAAA GACTGTGCAC   2040
CTTGCCAACC TTAACCTTTT GCCTCAGGGT GCCCAGGCCA CCTCTGAACT CCGCCAAGTG   2100
CTAACCAAAC CTCAGCAACA AATAAAGCAG GCAATAATCA ATGCAGCAGC CTCGCAACCC   2160
CCCAAAAAGG TGTCTCGAGT CCAGGTGGTG TCGTCCTTGC AGAGTTCTGT GGTGGAAGCT   2220
TTCAACAAGG TGCTGAGCAG TGTCAATCCA GTCCCTGTTT ACATCCCAAA CCTCAGTCCT   2280
CCCGCCAATG CAGGGATCAC GTTACCGACG CGTGGGTACA AGTGCTTGGA GTGTGGGGAC   2340
TCCTTTGCAC TTGAAAAGAG TCTGACCCAG CACTACGACA GACGGAGCGT GCGCATCGAA   2400
GTAACGTGCA ACCATTGTAC AAAGAACCTC GTTTTTTACA ACAAATGCAG CCTCCTTTCC   2460
CATGCCCGTG GGCATAAGGA GAAAGGGGTG GTAATGCAAT GCTCCCACTT AATTTTAAAG   2520
CCAGTCCCAG CAGATCAAAT GATAGTTTCT CCGTCAAGCA ATACTTCCAC TTCAACTTCC   2580
ACTCTTCAGA GCCCTGTGGG AGCTGGCACA CACACTGTCA CAAAAATTCA GTCTGGCATA   2640
ACTGGGACAG TCATATCGGC TCCTTCAAGC ACTCCCATCA CCCCAGCCAT GCCCCTAGAT   2700
GAAGACCCCT CCAAACTGTG TAGACATAGT CTAAAATGTT TGGAGTGTAA TGAAGTCTTC   2760
CAGGACGAGA CATCACTGGC TACACATTTC CAGCAGGCTG CAGATACGAG TGGACAAAAG   2820
```

```
ACTTGCACTA TCTGCCAGAT GCTGCTTCCT AACCAGTGCA GTTATGCATC ACACCAGAGA  2880
ATCCATCAGC ACAAATCTCC CTACACCTGC CCTGAGTGTG GGGCCATCTG CAGGTCGGTG  2940
CACTTCCAGA CCCACGTCAC CAAGAACTGT CTGCACTACA CGAGGAGAGT TGGTTTTCGA  3000
TGTGTGCATT GCAATGTTGT GTACTCTGAT GTGGCTGCTC TGAAGTCTCA CATTCAAGGT  3060
TCTCACTGTG AAGTCTTCTA CAAGTGTCCT ATTTGTCCAA TGGCCGTTTAA GTCTGCCCCA  3120
AGCACACATT CCCACGCCTA CACACAGCAT CCTGGCATCA AGATAGGAGA ACCAAAAATA  3180
ATATATAAGT GTTCCATGTG CGACACTGTG TTCACCCTGC AAACCTTGCT GTATCGCCAC  3240
TTTGACCAAC ACATTGAAAA CCAGAAGGTG TCTGTTTTCA AGTGTCCAGA CTGTTCTCTT  3300
TTATATGCAC AGAAGCAACT TATGATGGAC CATATCAAGT CTATGCATGG AACATTGAAA  3360
AGTATTGAAG GGCCTCCAAA CTTGGGTATA AACTTGCCTT TGAGCATTAA GCCTGCAACT  3420
CAAAATTCAG CAAATCAGAA CAAAGAGGAC ACCAAATCCA TGAATGGGAA AGAGAAATTG  3480
GAAAAGAAAT CTCCATCTCC TGTGAAAAAA TCAATGGAAA CCAAGAAAGT GGCCAGTCCT  3540
GGGTGGACGT GTTGGGAGTG TGACTGCCTG TTCATGCAGA GAGATGTGTA CATATCCCAC  3600
GTGAGGAAGG AGCACGGGAA GCAAATGAAG AAACACCCCT GCCGCCAGTG TGACAAGTCT  3660
TTCAGCTCGT CCCACAGCCT GTGCCGGCAC AACCGGATCA AGCACAAAGG CATCAGGAAA  3720
GTGTACGCCT GCTCGCACTG CCCAGACTCC AGACGTACCT TTACCAAACG TTTGATGCTG  3780
GAGAAGCACG TCCAGCTGAT GCATGGCATC AAGGACCCTG ACCTGAAAGA AATGACAGAT  3840
GCCACCAATG AGGAGGAAAC AGAAATAAAA GAAGACACTA AGGTCCCCAG TCCCAAGCGG  3900
AAGTTGGAAG AACCAGTTCT GGAGTTCAGG CCTCCCCGAG GAGCAATCAC TCAACCACTG  3960
AAAAAGCTGA AAATCAATGT TTTTAAGGTT CACAAGTGTG CCGTGTGTGG CTTCACCACC  4020
GAAAACCTGC TGCAATTCCA CGAACACATC CCTCAGCACA AATCGGATGG TTCTTCCTAC  4080
CAGTGCCGGG AGTGTGGCCT CTGCTACACG TCTCACGTCT CTCTGTCCAG GCACCTCTTC  4140
ATCGTACACA AGTTAAAGGA ACCTCAGCCA GTGTCCAAGC AAAATGGGGC TGGGGAAGAT  4200
AACCAACAGG AGAACAAACC CAGCCACGAG GATGAATCCC CTGATGGCGC CGTGTCAGAC  4260
AGAAAGTGCA AAGTGTGCGC AAAAACTTTT GAAACTGAAG CTGCCTTAAA TACTCACATG  4320
CGGACACACG GCATGGCCTT CATCAAATCC AAAAGGATGA GCTCAGCCGA GAAATAGCCA  4380
CAGATGCTCC ATGAGGAAAA TCCCTGTCCA CATTGGAATA AAAAAGACAT TTTTGTTACA  4440
AAGTTTGCAG TATAATAGAG TTAACAGTAC TGTCTAGGCT GTTGCAATAT ATTCTCTTTC  4500
AATGTACCTT CCTTCACCTC GTCGTATATA TCCTCGATAA GTATTAAAAC AGTATTTGAG  4560
TTTAAAAGAG TTTGTATATA TTTAAATGAA TAACTTTTTA TACTCTTTGT TACATGTTTG  4620
TATCAGTATT TAGTGGAAAA CCATTTGAGT TGTTTTGGGT TAGAATTTTT CTTTTTGTAC  4680
TGTTTCTTTA AAACAGAGTT CTTAGTAACA GGGGCAGTTC CTGAATTCAA ATAAACCATT  4740
TTGTATGTTT GGATTTTGAA TGGGTTAACT AATTACAGGC TAAAATAATG CCTTTTTTAG  4800
TGTTTTTAAT TTTTAGAATT CACTACATAA ATTGTAAGTA ATTGTGGGTC TCAAAAACAC  4860
TAGGAACTTT TAAGTGTCTT AGCACTTCCT CGATGTGCCT GCCCTGAGGG AGTGAGTTCA  4920
CATTTGAGAC AACTGCACTC CAGTGTGGAC GTGCCTTTGT CTTCAGGCCA TGCCGAAGGG  4980
TGTTTAAAGC AGTCTTGCAG GTCGCTCCTT TCCCAGCCGT GGATAAAAAC TGAAGCTAGG  5040
AATCTAATAA GGAATGCTGA TTTCCTCAGT TCCATTTTGA GGAATGGGGA AGGCTATTCT  5100
AAAGAAAAAA ATGGGATTTG TTTTCTCGGC AGATCTGCAA GGCTGGCTTT AAGAGCACAA  5160
GGAGGGAAAG TAACGAAAGG GCTGGACTAC TATAAAAGTT ACAAATACGT AGTTAGACCA  5220
ATAGATTTAT ATAGTCAGGT TTTTGTCATG TAATTTATTA ACTAACTATT ACAGAAACAC  5280
AGCTAAGAAT ATCAAGTATT TCTCTGGCTC TTGACAGAAA AAAATCAGTT GACTTAACCC  5340
TTTGCTGTCA AAAGAGTTGG CGTTTCCTGT TCTGGGTGCT ACTGCCAAAC GTTATGGTAC  5400
TTAGAGTCGG GATGCACAAC TTCAACCACC GACTTATCAA TGCAGCCGCC TGTGTATTGC  5460
AATTGGCCGT TACCTTAAGC ACTGAGCCAC CCGGGTTTAG TTCAGCCATT TCAAGAAGTA  5520
TATTTAACGT CGGTAGTTCT GCTTTATTAA AATGCAGCAG AGGTACTCTT CTGTCCCTTC  5580
CGTTTATAGT TCTCTGAGAG AGTTCTATTT TTTGGTTTTG TTTTGTGTTT TCTTTTGCAT  5640
TTTGTATCTT GTATTTATCC CTGAACATGT TTTGTACCTT TTTTTTTTTT TTTTTTTTAA  5700
GAAAAGGAAT TCTTTTGTGT ATATATAGAT ACTTGCATGA TATACTGTAG TCAATGTTCG  5760
GTTCCTCAAA AGGTCTTGCT GCTGTCAGGT GTTATGCACT CCATCCATCA TAACTGTATG  5820
AAACACATTT CATATGTAAA TAAACGTGGG ACATTTG
```

```
MKTPDFDDLL AAFDIPDMVD PKAAIESGHD DHESHMKQNA HGEDDSHAPS SSDVGVSVIV   60
KNVRNIDSSE GGEKDGHNPT GNGLHNGFLT ASSLDSYSKD GAKSLKGDVP ASEVTLKDST  120
FSQFSPISSA EEFDDDEKIE VDDPPDKEDM RSSFRSNVLT GSAPQQDYDK LKALGGENSS  180
KTGLSTSGNV EKNKAVKRET EASSINLSVY EPFKVRKAED KLKESSDKVL ENRVLDGKLS  240
SEKNDTSLPS VAPSKTKSSS KLSSCIAAIA ALSAKKAASD SCKEPVANSR ESSPLPKEVN  300
DSPRAADKSP ESQNLIDGTK KPSLKQPDSP RSISSENSSK GSPSSPAGST PAIPKVRIKT  360
IKTSSGEIKR TVTRVLPEVD LDSGKKPSEQ TASVMASVTS LLSSPASAAV LSSPPRAPLQ  420
SAVVTNAVSP AELTPKQVTI KPVATAFLPV SAVKTAGSQV INLKLANNTT VKATVISAAS  480
VQSASSAIIK AANAIQQQTV VVPASSLANA KLVPKTVHLA NLNLLPQGAQ ATSELRQVLT  540
KPQQQIKQAI INAAASQPPK KVSRVQVVSS LQSSVVEAFN KVLSSVNPVP VYIPNLSPPA  600
NAGITLPTRG YKCLECGDSF ALEKSLTQHY DRRSVRIEVT CNHCTKNLVF YNKCSLLSHA  660
RGHKEKGVVM QCSHLILKPV PADQMIVSPS SNTSTSTSTL QSPVGAGTHT VTKIQSGITG  720
TVISAPSSTP ITPAMPLDED PSKLCRHSLK CLECNEVFQD ETSLATHFQQ AADTSGQKTC  780
TICQMLLPNQ CSYASHQRIH QHKSPYTCPE CGAICRSVHF QTHVTKNCLH YTRRVGFRCV  840
HCNVVYSDVA ALKSHIQGSH CEVFYKCPIC PMAFKSAPST HSHAYTQHPG IKIGEPKIIY  900
KCSMCDTVFT LQTLLYRHFD QHIENQKVSV FKCPDCSLLY AQKQLMMDHI KSMHGTLKSI  960
EGPPNLGINL PLSIKPATQN SANQNKEDTK SMNGKEKLEK KSPSPVKKSM ETKKVASPGW 1020
TCWECDCLFM QRDVYISHVR KEHGKQMKKH PCRQCDKSFS SSHSLCRHNR IKHKGIRKVY 1080
ACSHCPDSRR TFTKRLMLEK HVQLMHGIKD PDLKEMTDAT NEEETEIKED TKVPSPKRKL 1140
EEPVLEFRPP RGAITQPLKK LKINVFKVHK CAVCGFTTEN LLQFHEHIPQ HKSDGSSYQC 1200
RECGLCYTSH VSLSRHLFIV HKLKEPQPVS KQNGAGEDNQ QENKPSHEDE SPDGAVSDRK 1260
CKVCAKTFET EAALNTHMRT HGMAFIKSKR MSSAEK
```

SEQ ID NO:256 PBM1 DNA sequence
Nucleic Acid Accession#: AF111847
Coding sequence: 58-1608 (underlined sequence corresponds to start and stop codon)

```
       1          11         21         31         41         51
       |          |          |          |          |          |
       TTTTCGTCGA CTCTTACCGG TTGGCTGGGC CAGCTGCGCC GCGGCTCACA GCTGACGATG    60
       GGGGACCCCA GCAAGCAGGA CATCTTGACC ATCTTCAAGC GCCTCCGCTC GGTGCCCACT   120
       AACAAGGTGT GTTTTGATTG TGGTGCCAAA AATCCCAGCT GGGCAAGCAT AACCTATGGA   180
       GTGTTCCTTT GCATTGATTG CTCAGGGTCC CACCGGTCAC TTGGTGTTCA CTTGAGTTTT   240
       ATTCGATCTA CAGAGTTGGA TTCCAACTGG TCATGGTTTC AGTTGCGATG CATGCAAGTC   300
       GGAGGAAACG CTAGTGCATC TTCCTTTTTT CATCAACATG GGTGTTCCAC CAATGACACC   360
       AATGCCAAGT ACAACAGTCG TGCTGCTCAG CTCTATAGGG AGAAAATCAA ATCGCTCGCC   420
       TCTCAAGCAA CACGGAAGCA TGGCACTGAT CTGTGGCTTG ATAGTTGTGT GGTTCCACCT   480
       TTGTCCCCTC CACCAAAGGA GGAAGATTTT TTTGCCTCTC ACGTTTCTCC TGAGGTGAGT   540
       GACACAGCGT GGGCATCAGC AATAGCAGAA CCATCTTCTT TAACATCAAG GCCTGTGGAA   600
       ACCACTTTGG AAAATAATGA AGGTGGACAA GAGCAAGGAC CAAGTGTGGA AGGTCTTAAT   660
       GTACCAACAA AGGCTACTTT AGAGGTATCC TCTATCATAA AAAAGAAACC AAATCAAGCT   720
       AAAAAAGGCC TTGGGGCCAA AAAAGGAAGT TTGGGAGCTC AGAAACTGGC AAACACATGC   780
       TTTAATGAAA TTGAAAAACA AGCTCAAGCT GCGGATAAAA TGAAGGAGCA GGAAGACCTG   840
       GCCAAGGTGG TATCTAAAGA AGAATCAATT GTTTCATCAT TACGATTAGC CTATAAGGAT   900
       CTTGAAATTC AAATGAAGAA AGACGAAAAG ATGAACATTA GTGGCAAAAA AAATGTTGAC   960
       TCAGACAGAC TCGGCATGGG ATTTGGAAAT TGCAGAAGTG TTATTTCACA TTCAGTGACT  1020
       TCAGATATGC AGACCATAGA GCAGGAATCA CCCATTATGG CAAAACCAAG AAAAAAGTAT  1080
       AATGATGACA GTGACGATTC ATATTTTACT TCCAGCTCAA GTTACTTTGA CGAGCCAGTG  1140
       GAGTTAAGGA GCAGTTCTTT CTCTAGCTGG GATGACAGTT CAGATTCCTA TTGGAAAAAA  1200
       GAGACCAGCA AAGATACTGA AACAGTTCTG AAAACCACAG GCTATTCAGA CAGACCTACT  1260
       GCTCGCCGCA AGCCAGATTA TGAGCCAGTT GAAAATACAG ATGAGGCCCA GAAGAAGTTT  1320
       GGCAATGTCA AGGCCATTTC ATCAGATATG TATTTTGGAA GACAATCCCA GGCTGATTAT  1380
       GAGACCAGGG CCCGCCTAGA GAGGCTGTCG GCAAGTTCCT CCATAAGCTC GGCTGATCTG  1440
       TTCGAGGAGC CGAGGAAGCA GCCAGCAGGG AACTACAGCC TGTCCAGTGT GCTGCCCAAC  1500
       GCCCCCGACA TGGCGCAGTT CAAGCAGGGA GTGAGATCGG TTGCTGGAAA ACTCTCCGTC  1560
       TTTGCTAATG GAGTCGTGAC TTCAATTCAG GATCGCTACG GTTCTTAATA CTGAAGTCAT  1620
       GATGTGTATT TCCTGGAGAA ATTCCTCTTT AAATGAACAA GTAACCACAT CTCAGGCGGC  1680
       AGTGAAGTCC AGATAGTTTT GCAGATTGTT TTGCTACTTT TTCATATGGT ATATGTTTCT  1740
       GATTTTTAAT ATTTCTTTTG AGAAATTCTG AGTTCTGATG TAGGAGCTTT CCTGTGATTT  1800
       CTGTTTCACG TTCCTTCCTG TCACACCCTC CTTTGGCGTC TCTGTGTATA TCCTTGCTTT  1860
       ATTTTCTTGG AACCTTTGAT TTCAACACTG AGGGCCTGGA GACCTCGGCT CCTCCTGCTC  1920
       CTGAACCAGG AGGCTTCATG TGGGGGAGGA GGAGAGGTCT CCATGTGACA CATGGGCTCA  1980
       GGGCTGCCAG AATCAGCGGA TGCTGGATGG GCCTGCAGAA ACAACACTCA CCACACACAC  2040
       TTCCTTCAAA AGACCAAAAG TGACTGGTGT CTCGTGTGAC AGATTGCTTC ATTTATGTTT  2100
       CTACATAGTA AGGTGACTGC CAAATAATAT TTGAAGTCAT CTGTCTCTTT GTAAATTATT  2160
       TTATATGACC TATAAATTTA AAAATGTTTT TCAGTGAGTG CTTTTAACAA ACTTAAGCTT  2220
       CTGCCCTGCC AAGGGAATTA ATGTTATCTT GTGAAAGGTG TTGCTGTTTG AATTGATGAG  2280
       AAATGGAAGA TGAGAACTCC CTAAGAGTTC TCATAATAAA TCATCTCATC ACAAATCAAT  2340
       ACGGTATACA GAGTTAAAGT GGAATGAGGT AAGAAGATAC AGCTACAGAA AATAGTTGCG  2400
       TGTATGGGAG AACAGTCATT GTAATTGGGT AGTTTTGTTA ATAAATATTT TTAAATCTTG  2460
       CTTTTCAGAA ATTACCGAAT GTGTATAAAC AAATAAAGAA AAATAATTTA GCTGTGTTTT  2520
       AGACAGCATT AGAATATATT GTTCAGCACA GTAAAATATA TTTGAAATTT GATAAGCCAA  2580
       AAATGTGGTT TTGAATGAAT ATTTTGTGAA TCTTTCTTAA AAGCTCAAAT TTGTAGACTT  2640
       CTAAATAGAA TAAACACTTG CAGCAGAAAA AAAAAAAAAA AAAAAAAAA AAAAAAAAAA  2700
       AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA  2760
       AAAAAAAAAA AAAAAAAAAA AAAAAAAA
```

SEQ ID NO:257 PBM1 Protein sequence:
PBM1 Protein sequence: CAB76901

```
MGDPSKQDIL TIFKRLRSVP TNKVCFDCGA KNPSWASITY GVFLCIDCSG SHRSLGVHLS   60
FIRSTELDSN WSWFQLRCMQ VGGNASASSF FHQHGCSTND TNAKYNSRAA QLYREKIKSL  120
ASQATRKHGT DLWLDSCVVP PLSPPPKEED FFASHVSPEV SDTAWASAIA EPSSLTSRPV  180
ETTLENNEGG QEQGPSVEGL NVPTKATLEV SSIIKKKPNQ AKKGLGAKKG SLGAQKLANT  240
CFNEIEKQAQ AADKMKEQED LAKVVSKEES IVSSLRLAYK DLEIQMKKDE KMNISGKKNV  300
DSDRLGMGFG NCRSVISHSV TSDMQTIEQE SPIMAKPRKK YNDDSDDSYF TSSSSYFDEP  360
VELRSSSFSS WDDSSDSYWK KETSKDTETV LKTTGYSDRP TARRKPDYEP VENTDEAQKK  420
FGNVKAISSD MYFGRQSQAD YETRARLERL SASSSSISSAD LFEEPRKQPA GNYSLSSVLP  480
NAPDMAQFKQ GVRSVAGKLS VFANGVVTSI QDRYGS
```

SEQ ID NO:258 PBM4 DNA sequence
Nucleic Acid Accession#: D30891
Coding sequence: 1-4032 (underlined sequence corresponds to start and stop codon)

```
ATGGATACTG TCATGAAGCA GACACATGCT GACACACCTG TTGATCATTG TCTATCTGGC   60
ATAAGAAAGT GTAGCAGCAC CTTTAAGCTT AAAAGTGAAG TCAACAAGCA TGAAACAGCC  120
CTTGAAATGC AGAATCCAAA TTTGAACAAT AAAGAATGTT GTTTCACCTT TACGTTGAAT  180
GGAAACTCCA GAAAATTAGA CCGTAGTGTG TTTACAGCAT ATGGTAAACC CAGCGAGAGT  240
ATCTACTCAG CCCTGAGTGC TAATGACTAT TTCAGTGAAA GGATAAAGAA TCAGTTTAAT  300
AAGAACATTA TTGTTTATGA AGAAAAGACA ATAGATGGAC ATATAAATTT AGGAATGCCT  360
CTCAAGTGCC TGCCTAGTGA TTCTTCATTTT AAAATTACAT TTGGTCAAAG AAAGAGTAGC  420
```

```
AAAGAAGATG GACACATATT ACGCCAATGT GAAAATCCAA ACATGGAATG CATTCTTTTT  480
CATGTTGTTG CTATAGGAAG GACAAGAAAG AAGATTGTTA AGATCAACGA ACTTCATGAA  540
AAAGGAAGTA AACTTTGTAT TTATGCCTTG AAGGGTGAGA CTATTGAAGG AGCCTTATGC  600
AAGGATGGCC GTTTTCGGTC TGACATAGGT GAATTTGAAT GGAAACTAAA GGAAGGTCAT  660
AAGAAAATTT ATGGAAAACA GTCCATGGTG GATGAAGTAT CTGGAAAAGT CTTAGAAATG  720
GACATTTCAA AAAAAAAAGC ATTACAACAG AAAGATATCC ATAAAAAAAT TAAACAGAAT  780
GAAAGTGCCA CTGATGAAAT TAATCACCAG AGTCTGATAC AGTCTAAGAA AAAAGTCCAC  840
AAACCAAAGA AAGATGGAGA GACCAAAGAT GTAGAACACA GCAGAGAGCA AATTCTCCCA  900
CCTCAGGATC TAAGCCATTA TATTAAAGAT AAAACTCGCC AGACAATTCC CAGGATTAGA  960
AATTATTACT TTTGTAGTTT GCCCCGAAAA TATAGGCAAA TAAACTCACA AGTTAGACGG  1020
AGGCCGCATC TGGGTAGGCG GTATGCTATT AATCTGGATG TCCAAAAGGA GGCAATTAAT  1080
CTCTTAAAGA ATTATCAAAC GTTGAATGAA GCCATAATGC ATCAGTATCC GAATTTTAAA  1140
GAGGAGGCAC AGTGGGTAAG AAAATATTTT CGGGAAGAAC AAAAGAGAAT GAATCTTTCA  1200
CCAGCTAAGC AATTCAACAT ATATAAAAAG GACTTCGGAA AAATGACTGC AAATTCTGTT  1260
TCAGTTGCAA CCTGCGAACA GCTTACATAT TATAGCAAGT CAGTTGGGTT CATGCAATGG  1320
GACAATAATG GAAACACAGG TAATGCTACT TGCTTTGTCT TCAATGGTGG TTATATTTTC  1380
ACCTGTCGAC ATGTTGTACA TCTTATGGTG GGTAAAAACA CACATCCAAG TTTGTGGCCA  1440
GATATAATTA GCAAATGTGC GAAGGTAACC TTCACTTATA CAGAGTTCTG CCCTACTCCT  1500
GACAATTGGT TTTCCATTGA GCCATGGCTT AAAGTGTCCA ATGAAAATCT AGATTATGCC  1560
ATTTTAAAAC TAAAAGAAAA TGGAAATGCG TTTCCTCCAG GACTATGGCG ACAGATTTCT  1620
CCTCAACCAT CTACTGGTTT GATTTATTTA ATTGGTCATC CTGAAGGCCA GATCAAGAAA  1680
ATAGATGGTT GTACTGTGAT TCCTCTAAAC GAACGATTGA AAAAATATCC AAACGATTGT  1740
CAAGATGGGT TGGTAGATCT CTATGATACC ACCAGTAATG TATACTGTAT GTTTACCCAA  1800
AGAAGTTTCC TATCAGAGGT TTGGAACACA CACACGCTTA GTTATGATAC TTGTTTCTCT  1860
GATGGGTCCT CAGGCTCCCC AGTGTTTAAT GCATCTGGCA AATTGGTTGC TTTGCATACC  1920
TTTGGGCTTT TTTATCAACG AGGATTTAAT GTGCATGCCC TTATTGAATT TGGTTATTCT  1980
ATGGATTCTA TTCTTTGTGA TATTAAAAAG ACAAATGAGA GCTTGTATAA ATCATTAAAT  2040
GATGAGAAAC TTGAGACCTA CGATGAAGAG AAAGCCCGGC CCAGGCCAGC CTACCGGCGA  2100
CTAGGATGCT TTCGCTTTCG CTCTCGCTTT CCAATACTCG GGACTGGGGA AACCGGGAGA  2160
ATAGAAGCAG GCAAGGACCG CCGTGGGCAC GGGGTCAGTG AGACAGGGTC CTGCTCGCGG  2220
CGTCAAGGAG GAGCGCTGTG GGTGTCCCCA GCGCAGCCAA TCGGCTTCCG AAGTAGCTGG  2280
AGCTCTGGAG CCTTTGCTTC CTCAAATACG AGCGGGAACT GCGTTGAGCG CTGGATTCCA  2340
GGCCGAGTGC TGGCGAGGCG CGCAGTCTCT AAAGAGCAAC AGAATAATTG CAGTACTTCT  2400
CTAATGAGGA TGGAGTCTAG AGGAGACCCA AGAGCCACAA CTAATACCCA GGCTCAAAGA  2460
TTCCATTCAC CTAAGAAAAA TCCAGAAGAC CAGACCATGC CCCAAAATAG GACAATATAT  2520
GTTACCTTGA AGGCTGTCAA AAAAGAGATA GAAACTCACC AAGGCCAAGA AATGCTTGTG  2580
CGTGGCACAG AAGGAATCAA AGAGTACATA AACCTTGGAA TGCCCCTCAG TTGTTTCCCT  2640
GAAGGTGGCC AGGTGGTCAT TACATTTTCC CAAAGTAAAA GTAAGCAGAA GGAAGATAAC  2700
CACATATTTG GCAGGCAGGA CAAAGCATCG ACTGAATGTG TCAAATTTTA CATTCATGCA  2760
ATTGGAATTG GGAAGTGTAA AAGAAGGATT GTTAAATGTG GGAAGCTTCA CAAAAAGGGG  2820
CGCAAACTCT GTGTTTATGC TTTCAAAGGA GAAACCATCA AGGATGCACT GTGCAAGGAT  2880
GGCAGATTTC TTTCCTTTCT GGAGAATGAT GATTGGAAAC TCATTGAAAA CAATGACACC  2940
ATTTTAGAAA GCACCCAGCC AGTTGATGAA TTAGAAGGCA GATACTTTCA GGTTGAGGTT  3000
GAGAAAAGAA TGGTCCCCAG TGCAGCAGCT TCTCAGAATC CTGAGTCAGA GAAAAGAAAC  3060
ACCTGTGTGT TGAGAGAACA AATCGTGGCT CAGTACCCCA GTTTGAAAAG AGAAAGTGAA  3120
AAAATCATTG AAAACTTCAA GAAAAAAATG AAAGTAAAAA ATGGGGAAAC ATTATTTGAA  3180
TTGCATAGAA CAACGTTTGG GAAAGTAACA AAAAATTCTT CTTCGATTAA AGTAGTGAAA  3240
CTTCTTGTAC GTCTCAGTGA CTCAGTTGGG TACTTATTCT GGGACAGTGC AACTACGGGT  3300
TACGCCACCT GCTTTGTTTT TAAAGGATTG TTCATTTTAA CTTGTCGGCA TGTAATAGAT  3360
AGCATTGTGG GAGACGGAAT AGAGCCAAGT AAGTGGGCAA CCATAATTGG TCAATGTGTA  3420
AGGGTGACAT TTGGTTATGA AGAGCTAAAA GACAAGGAAA CAAACTACTT TTTTGTTGAA  3480
CCTTGGTTTG AGATACATAA TGAAGAGCTT GACTATGCTG TCCTGAAACT GAAGGAAAAT  3540
GGACAACAAG TACCTATGGA ACTATATAAT GGAATTACTC CTGTGCCACT TAGTGGGTTG  3600
ATACATATTA TTGGCCATCC ATATGGAGAA AAAAAGCAGA TTGATGCTTG TGCTGTGATC  3660
CCTCAGGGTC AGCGAGCAAA GAAATGTCAG GAACGTGTTC AGTCTAAAAA AGCAGAAAGT  3720
CCAGAGTATG TCCATATGTA TACTCAAAGA AGTTTCCAGA AAATAGTTTCA CAACCCTGAT  3780
GTGATTACCT ATGACACTGA ATTTTTCTTT GGGGCTTCCG GCTCCCCTGT GTTTGATTCA  3840
AAAGGTTCAT TGGTGGCCAT GCATGCTGCT GGCTTTGCTT ATACTTACCA AAATGAGACT  3900
CGTAGTATCA TTGAGTTTGG CTCTACCATG GAATCCATCC TCCTTGATAT TAAGCAAAGA  3960
CATAAACCAT GGTATGAAGA AGTATTTGTA AATCAGCAGG ATGTAGAAAT GATGAGTGAT  4020
GAGGACTTGT GAGAATTCAG TCTACTGGAT TTAAGGGAAT GGCTTATGGA GTTGTTATTT  4080
CGTAGGCATT GAAAATGGTT TTCTAAACTC CAAAATGGTC ATCTTATCAA TAATAATAAT  4140
ATTGACCATT TCCTATCTGC CAGGCATTTT TCTAAGCACA TGAAGAAATT AGTCCTAACA  4200
ACACTATGAG ATGGACTATA ACTTGCCCAA ATTTTTTTTT TTTTTGAGAC TGAGTCTCAC  4260
TCTGTCGCCT GGGCTGGAGT ACAGTGGTGC GATCTCAGCT CACTGCAACT TCCACCTCCC  4320
AGGTTCAAGC GATTCTTATG CCTCAGTCTC CTGAGCAGCT GGGATTACAG GCAAACGCCA  4380
CCACACCCAG CTAAATTTTT TTTTTTTTTT TGTATTTTTA GTAGAGACAG GGTTTCACCA  4440
TGTTGGTCAG GCGGGTCTCG AACTCCTGAC CTCGTGATCC ACCTGCCTCG GCCTTCCAAA  4500
GTGCTGGGAT TACAAGTTTG AGCCACTGCA CCTGGCTAAC TTGCCCTATT TTAAAGTCAA  4560
GCAATGGGAA GAATAACAAG ATTATATAGT AATCAGTTTC ATGACACTAA AAGTCATATA  4620
GTCATAGGGT TTTTTCATCT TTCATATCTT TGCCTAAATT CATTTGCTAC AGTGCAGGAA  4680
CCAAAACTTG TTCATCTCAT GATTCCCTAC ATCTGACATA AGGAAAGTAA GTGCTCAGAA  4740
AAATGTGCAG GTCAATAAGT TGCAAAAGTT GGGGCTGCAA TTAATGCTAA CATAAGAGCT  4800
AAATGCTTGA TTAGAAATGA TCTCAAAACC TTTTAGAATT TCCAAAATCT TCATATTACT  4860
GAAACTGTCG GAATATATGG GTCCTGAAAT TCAGAAGATG ATAGTCACTC TTCCCATATT  4920
TATAGGCTAT TAAGGCAAGG GATATCTTAA ACATCATATT ACTTTATTTA GATTTCTACT  4980
ACTCCAATTA TTAATGTTAT GTATTTCTCA TTGTTTTACT TCTTCATGGT ATTATGAAGA  5040
CTATATAGAT GATTCAACCA AGCCTGCAAA TCTCCCTCTT GTGGAATTCC ACTGGACCCA  5100
ATCTGTTTTC CATTTCCATT GCAATACTAC TAAAGCCATA CAATATCAAG CACCCTCCCT  5160
```

410

CTAGGTCCAG GGACTATCAC AGAAGAAGCA GGCATGTAAG ATTTTAAGGA CTGGTTTCGA 5220
GGGGTCGAGT GTAGGAAAAC AGCCTGTTGC ATTGTAAGAG TGATGTCACC TTGAAGAGCA 5280
GCTGGCATGA TGACTGCTGT TTGACTCCTG CATACCAAGA TATTCTGCAG CAATGTCTTT 5340
AAACAGTGCC GGTAGTACAG ATAACCCCTC ATAAAGATGC TTATCTAACC TCCCCAGTGT 5400
TCAGGTGTTT CACAAGAAAG TCTGAGATAT GACTAGCTAC ACGTTTTGCC AAAAATGCTT 5460
GTTATATAAA GGGTACTTTT GGGAGGGTGA GTGCCGCCAT TTAGTGGCTG CTAGAAACAT 5520
TGCTTCTGTT TGTAAGTTCC TATTAAATGT TCTTTCTGAG AAAAAAAAAA A

<u>SEQ ID NO:259 PBM4 Protein sequence:</u>
PBM4 Protein sequence: BAB67788

MDTVMKQTHA DTPVDHCLSG IRKCSSTFKL KSEVNKHETA LEMQNPNLNN KECCFTFTLN   60
GNSRKLDRSV FTAYGKPSES IYSALSANDY FSERIKNQFN KNIIVYEEKT IDGHINLGMP   120
LKCLPSDSHF KITFGQRKSS KEDGHILRQC ENPNMECILF HVVAIGRTRK KIVKINELHE   180
KGSKLCIYAL KGETIEGALC KDGRFRSDIG EFEWKLKEGH KKIYGKQSMV DEVSGKVLEM   240
DISKKKALQQ KDIHKKIKQN ESATDEINHQ SLIQSKKKVH KPKKDGETKD VEHSREQILP   300
PQDLSHYIKD KTRQTIPRIR NYYFCSLPRK YRQINSQVRR RPHLGRRYAI NLDVQKEAIN   360
LLKNYQTLNE AIMHQYPNFK EEAQWVRKYF REEQKRMNLS PAKQFNIYKK DFGKMTANSV   420
SVATCEQLTY YSKSVGFMQW DNNGNTGNAT CFVFNGGYIF TCRHVVHLMV GKNTHPSLWP   480
DIISKCAKVT FTYTEFCPTP DNWFSIEPWL KVSNENLDYA ILKLKENGNA FPPGLWRQIS   540
PQPSTGLIYL IGHPEGQIKK IDGCTVIPLN ERLKKYPNDC QDGLVDLYDT TSNVYCMFTQ   600
RSFLSEVWNT HTLSYDTCFS DGSSGSPVFN ASGKLVALHT FGLFYQRGFN VHALIEFGYS   660
MDSILCDIKK TNESLYKSLN DEKLETYDEE KARPRPAYRR LGCFRFRSRF PILGTGETGR   720
IEAGKDRRGH GVSETGSCSR RQGGALWVSP AQPIGFRSSW SSGAFASSNT SGNCVERWIP   780
GRVLARRAVS KEQQNNCSTS LMRMESRGDP RATTNTQAQR FHSPKKNPED QTMPQNRTIY   840
VTLKAVRKEI ETHQGQEMLV RGTEGIKEYI NLGMPLSCFP EGGQVVITFS QSKSKQKEDN   900
HIFGRQDKAS TECVKFYIHA IGIGKCKRRI VKCGKLHKKG RKLCVYAFKG ETIKDALCKD   960
GRFLSFLEND DWKLIENNDT ILESTQPVDE LEGRYFQVEV EKRMVPSAAA SQNPESEKRN   1020
TCVLREQIVA QYPSLKRESE KIIENFKKKM KVKNGETLFE LHRTTFGKVT KNSSSIKVVK   1080
LLVRLSDSVG YLFWDSATTG YATCFVFKGL FILTCRHVID SIVGDGIEPS KWATIIGQCV   1140
RVTFGYEELK DKETNYFFVE PWFEIHNEEL DYAVLKLKEN GQQVPMELYN GITPVPLSGL   1200
IHIIGHPYGE KKQIDACAVI PQGQRAKKCQ ERVQSKKAES PEYVHMYTQR SFQKIVHNPD   1260
VITYDTEFFF GASGSPVFDS KGSLVAMHAA GFAYTYQNET RSIIEFGSTM ESILLDIKQR   1320
HKPWYEEVFV NQQDVEMMSD EDL

<u>SEQ ID NO:260 PBQ1 DNA sequence</u>
Nucleic Acid Accession#:   NM_015642
Coding sequence: 489-2489 (underlined sequence corresponds to start and stop codon)

```
        1         11        21        31        41        51
        |         |         |         |         |         |
    ACATTTCAAA AAAAATACAT AGACTGATGT TTCAGACTTG TGCAGCATAA GCCTACAGGG    60
    TACGAAGAAT GAACTCTGAG AATGTTTGGA GAATGTTTCA TCATTACTAA CAGGATATTC    120
    CTCATGACAT TGCTGTCTGA TCTTTGACCA TCAGTCTGTG ACCTGCCCCT TCTCTTTACA    180
    TGCAGCCGCT CTCTGCTCCC TGCCCCAATG AACATCTGCA CTAGGCCCAA GCCTTGGAGT    240
    AATTTACCTG AAGAGTGACA CCATTGATTT TGAAACTACT GAAGAAACCC AAGACAGCTG    300
    AAAAACCAGAA GGCATCTGAG GAGAATGAGA TTACTCAGCC GGGTGGATCC AGCGCCAAGC    360
    CGGGCCTTCC CTGCCTGAAC TTTGAAGCTG TTTTGTCTCC AGACCCAGCC CTCATCCACT    420
    CAACACATTC ACTGACAAAC TCTCACGCTC ACACCGGGTC ATCTGATTGT GACATCAGTT    480
    GCAAGGGGAT GACCGAGCGC ATTCACAGCA TCAACCTTCA CAACTTCAGC AATTCCGTGC    540
    TCGAGACCCT CAACGAGCAG CGCAACCGTG GCCACTTCTG TGACGTAACG GTGCGCATCC    600
    ACGGGAGCAT GCTGCGCGCA CACCGCTGCG TGCTGGCAGC CGGCAGCCCC TTCTTCCAGG    660
    ACAAACTGCT GCTTGGCTAC AGCGACATCG AGATCCCGTC GGTGGTGTCA GTGCAGTCAG    720
    TGCAAAAGCT CATTGACTTC ATGTACAGCG GCGTGCTACG GGTCTCGCAG TCGGAAGCTC    780
    TGCAGATCCT CACGGCCGCC AGCATCCTGC AGATCAAAAC AGTCATCGAC GAGTGCACGC    840
    GCATCGTGTC ACAGAACGTG GGCGATGTGT TCCCGGGGAT CCAGGACTCG GGCCAGGACA    900
    CGCCGCGGGG CACTCCCGAG TCAGGCACGT CAGGCCAGAG CAGCGACACG GAGTCGGGCT    960
    ACCTGCAGAG CCACCCACAG CACAGCGTGG ACAGGATCTA CTCGGCACTC TACGCGTGCT    1020
    CCATGCAGAA TGGCAGCGGC GAGCGCTCTT TTTACAGCGG CGCAGTGGTC AGCCACCACG    1080
    AGACTGCGCT CGGCCTGCCC CGCGACCACC ACATGGAAGA CCCCAGCTGG ATCACACGCA    1140
    TCCATGAGCG CTCGCAGCAG ATGGAGCGCT ACCTGTCCAC CACCCCCGAG ACCACGCACT    1200
    GCCGCAAGCA GCCCCGGCCT GTGCGCATCC AGACCCTAGT GGGCAACATC CACATCAAGC    1260
    AGGAGATGGA GGACGATTAC GACTACTACG GGCAGCAAAG GGTGCAGATC CTGGAACGCA    1320
    ACGAATCCGA GGAGTGCACG GAAGACACAG ACCAGGCCGA GGGCACCGAG AGTGAGCCCA    1380
    AAGGTGAAAG CTTCGACTCG GGCGTCAGCT CCTCCATAGG CACCGAGCCT GACTCGGTGG    1440
    AGCAGCAGTT TGGGCCTGGG GCGGCGCGGG ACAGCCAGGC TGAACCCACC CAACCCGAGC    1500
    AGGCTGCAGA AGCCCCCGCT GAGGGTGGTC CGCAGACAAA CCAGCTAGAA ACAGGTGCTT    1560
    CCTCTCCGGA GAGAAGCAAT GAAGTGGAGA TGGACAGCAC TGTTATCACT GTCAGCAACA    1620
    GCTCCGACAA GAGCGTCCTA CAACAGCCTT CGGTCAACAC GTCCATCGGG CAGCCATTGC    1680
    CAAGTACCCA GCTCTACTTA CGCCAGACAG AAACCCTCAC CAGCAACCTG AGGATGCCTC    1740
    TGACCTTGAC CAGCAACACG CAGGTCATTG GCACAGCTGG CAACACCTAC CTGCCAGCCC    1800
    TCTTCACTAC CCAGCCCGCG GGCAGTGGCC CCAAGCCTTT CCTCTTCAGC CTGCCACAGC    1860
    CCCTGGCAGG CCAGCAGACC CAGTTTGTGA CAGTGTCCCA GCCCGGTCTG TCGACCTTTA    1920
    CTGCACAGCT GCCAGCGCCA CAGCCCCTGG CCTCATCCGC AGGCCACAGC ACAGCCAGTG    1980
    GGCAAGGCGA AAAAAAGCCT TATGAGTGCA CTCTCTGCAA CAAGACTTTC ACCGCCAAAC    2040
    AGAACTACGT CAAGCACATG TTCGTACACA CAGGTGAGAA GCCCCACCAA TGCAGCATCT    2100
    GTTGGCGCTC CTTCTCCTTA AAGGATTACC TTATCAAGCA CATGGTGACA CACACAGGAG    2160
```

411

```
TGAGGGCATA CCAGTGTAGT ATCTGCAACA AGCGCTTCAC CCAGAAGAGC TCCCTCAACG  2220
TGCACATGCG CCTCCACCGG GGAGAGAAGT CCTACGAGTG CTACATCTGC AAAAAGAAGT  2280
TCTCTCACAA GACCCTCCTG GAGCGACACG TGGCCCTGCA CAGTGCCAGC AATGGGACCC  2340
CCCCTGCAGG CACACCCCCA GGTGCCCGCG CTGGCCCCCC AGGCGTGGTG GCCTGCACGG  2400
AGGGGACCAC TTACGTCTGC TCCGTCTGCC CAGCAAAGTT TGACCAAATC GAGCAGTTCA  2460
ACGACCACAT GAGGATGCAT GTGTCTGACG GATAAGTAGT ATCTTTCTCT CTTTCTTATG  2520
AACAAAACAA AACAACAACA AAAAACAAAC AAACAAAAAA GCTATGGCAC TAGAATTTAA  2580
GAAATGTTTT GGTTTCATTT TTACTTTCTG TTTTTGTTTT TGTTTCGTTT CATTTTGTAC  2640
TACATGAAGA ACTGTTTTTT GCCTGCTGGT ACATTACATT TCCGGAGGCT TGGGTGAATA  2700
ATAGTTTTCC CAGTCTCCCT CGGATGGTGG CCTTAAGGCC TGGTAGTGCT TCAAGAGGTC  2760
CACTGGTTGG ATCTCTAGCT ACTGGCCTCT AAATACAACC CTTCTTTACA AAAAAAAAAA  2820
AAAAAAAAA
```

SEQ ID NO:261 PBQ1 Protein sequence:
PBQ1 Protein sequence: NP_056457

```
MTERIHSINL HNFSNSVLET LNEQRNRGHF CDVTVRIHGS MLRAHRCVLA AGSPFFQDKL   60
LLGYSDIEIP SVVSVQSVQK LIDFMYSGVL RVSQSEALQI LTAASILQIK TVIDECTRIV  120
SQNVGDVFPG IQDSGQDTPR GTPESGTSGQ SSDTESGYLQ SHPQHSVDRI YSALYACSMQ  180
NGSGERSFYS GAVVSHHETA LGLPRDHHME DPSWITRIHE RSQQMERYLS TTPETTHCRK  240
QPRPVRIQTL VGNIHIKQEM EDDYDYYGQQ RVQILERNES EECTEDTDQA EGTESEPKGE  300
SFDSGVSSSI GTEPDSVEQQ FGPGAARDSQ AEPTQPEQAA EAPAEGGPQT NQLETGASSP  360
ERSNEVEMDS TVITVSNSSD KSVLQQPSVN TSIGQPLPST QLYLRQTETL TSNLRMPLTL  420
TSNTQVIGTA GNTYLPALFT TQPAGSGPKP FLFSLPQPLA GQQTQFVTVS QPGLSTFTAQ  480
LPAPQPLASS AGHSTASGQG EKKPYECTLC NKTFTAKQNY VKHMFVHTGE KPHQCSICWR  540
SFSLKDYLIK HMVTHTGVRA YQCSICNKRF TQKSSLNVHM RLHRGEKSYE CYICKKKFSH  600
KTLLERHVAL HSASNGTPPA GTPPGARAGP PGVVACTEGT TYVCSVCPAK FDQIEQFNDH  660
MRMHVSDG
```

SEQ ID NO: 262 PBQ6 DNA sequence
Nucleic Acid Accession#:   AI654187
Coding sequence: 1-912 (underlined sequence corresponds to start and stop codon)

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGGTGGAAG AGGAAACAGG CATATCTTAC ATGGTGGCAG ACAAGGGACA CCCTTCTACA   60
AACTCTACCA CTTCTGCGCC GTCGTTTCGA CCATATAAAA ACGACCTATG CGAACTGCGT  120
CGGAAAACTC CCTCACGATG TAAAACGAAG ATCAGGAGCA GATTTGAAGA ATTACAAAGT  180
GAATTGGTGC CAGTCAGCAT GTCAGAGACA GACCACATAG CCTCTACTTC CTCTGATAAA  240
AATGTTGGGA AAACACCTGA ATTAAAGGAA GACTCATGCA ACTTGTTTTC TGGCAATGAA  300
AGCAGCAAAT TAGAAAATGA GTCCAAACTA TTGTCATTAA ACACTGATAA AACTTTATGT  360
CAACCTAATG AGCATAATAA TCGAATTGAA GCCCAGGAAA ATTATATTCC AGATCATGGT  420
GGAGGTGAGG ATTCTTGTGC CAAAACAGAC ACAGGCTCAG AAAATTCTGA ACAAATAGCT  480
AATTTTCCTA GTGGAAATTT TGCTAAACAT ATTTCAAAAA CAAATGAAAC AGAACAGAAA  540
GTAACACAAA TATTGGTGGA ATTAAGGTCA TCTACATTTC CAGAATCAGC TAATGAAAAG  600
ACTTATTCAG AAAGCCCCTA TGATACAGAC TGCACCAAGA AATTTATTTC AAAAATAAAG  660
AGCGTTTCAG CATCAGAGGA TTTGTTGGAA GAAATAGAAT CTGAGCTCTT ATCTACGGAG  720
TTTGCAGAAC ATCGAGTACC AAATGGAATG AATAAGGGAG AACATGCATT AGTTCTGTTT  780
GAAAAGTGTG TGCAAGATAA ATATTTGCAG CAGGAACATA TCATAAAAAA GGCCAGACTT  840
GGTCTCTGTT ATTTGCCATC AAGAACCTCA ATTGACACGT TAATTCCGTT TATCCCAAAT  900
TTATATAGAT AA
```

SEQ ID NO:263 PBQ6 Protein sequence:
Protein Accession #:   NP_060170

```
MEPKEATGKE NMVTKKKNLA FLRSRLYMLE RRKTDTVVES SVSGDHSGTL RRSQSDRTEY   60
NQKLQEKMTP QGECSVAETL TPEEEHHMKR MMAKREKIIK ELIQTEKDYL NDLELCVREV  120
VQPLRNKKTD RLDVDSLFSN IESVHQISAK LLSLLEEATT DVEPAMQVIG EVFLQIKGPL  180
EDIYKIYCYH HDEAHSILES YEKEEELKEH LSHCIQSLK
```

SEQ ID NO:264 PBY7 DNA sequence
Nucleic Acid Accession#:   NM_014323
Coding sequence: 662-2725 (underlined sequence corresponds to start and stop codon)

```
1          11         21         31         41         51
|          |          |          |          |          |
GGGCCTACTC TGCCGCCGCC GCCGCCCGCC CGCTCCAGCC GCCGCCGCCG CCGCCACCGC   60
CCTCCAGGCT CCGGGACCCG GCCCGCGCCA CCGCCCCCGT GCGCGCCCCG CCGCCGCCGC  120
CTTCGCCTTC GCCTTTTGTT TCCTCCGCTC CGGCGCCCCC GCCCCGGCTC GCGCTTTGCA  180
GGGGACGCAG CGCGCGCCCC CAGCGGGCCC GGGAAAAGCC GCGGCGCGCG CGCGCGCCTG  240
CGCGGCGGAC CCCTCCTTCT CCTCCCCGCG TGCGCGTGCC CTTCTTGGCT GCGCGCCGGC  300
GCCGCCTGGC GGGCGGGAGG GGAGGTGGCA GGCGCGTTTG CAGGAGGGGC GCACCTCTTC  360
GCTCGCGCAC CCCCCCGGAA GGTAGACCGG GAAGGGGAGG CGGGCGGGCG GAGAGGAGAG  420
AGTGGCGGCG AGTCCAGCGA GGGCGGGGGT TGGCTATGTG GGGGGTGGTG CACCCCGCAG  480
TCTAGACAGT CTGATCCGGG CTGGGGGCGT GTACACTCGG CGCACCTGCG AGACTACAGA  540
GCCTCGGGCC GGCACGTGTG GGGAGTGTGG ACACGTCTGC TGCGCCCCGC TTCTCGCTGC  600
```

412

```
TGAGGGGAAG GGAGGGGGCG GGCAGGTGCA GCGGCCGGGC TAGTGGGAGG GGGCGGCGGC   660
CATGGAGCGG GTGAACGACG CTTCGTGCGG CCCGTCTGGC TGCTACACAT ACCAGGTGAG   720
CAGACACAGC ACGGAGATGC TGCACAACCT GAACCAGCAG CGCAAAAACG GCGGGCGCTT   780
CTGCGACGTG CTCTTGCGGG TAGGCGACGA GAGCTTCCCA GCGCACCGCG CCGTGCTGGC   840
CGCCTGCAGC GAGTACTTTG AGTCGGTGTT CAGCGCCCAG TTGGGCGACG GCGGAGCTGC   900
GGACGGGGGT CCGGCTGATG TAGGGGGCGC GACGGCAGCA CCAGGCGGCG GGGCCGGGGG   960
CAGCCGGGAG CTGGAGATGC ACACTATCAG CTCCAAGGTA TTTGGGGACA TTCTGGACTT  1020
CGCCTACACT TCCCGCATCG TGGTGCGCTT GGAGAGCTTT CCCGAACTCA TGACGGCCGC  1080
CAAGTTCCTG CTGATGAGGT CGGTTATCGA GATCTGCCAG GAAGTCATCA AACAGTCCAA  1140
CGTACAGATC CTGGTACCCC CTGCCCGCGC CGATATAATG CTCTTTCGCC CCCCTGGGAC  1200
CTCGGACTTG GGCTTCCCTT TGGACATGAC CAACGGGGCA GCCTTGGCAG CCAACAGCAA  1260
TGGCATCGCC GGCAGCATGC AGCCAGAGGA GGAGGCAGCT CGGGCGGCTG GTGCAGCCAT  1320
TGCAGGCCAA GCCTCTTTGC CTGTGTTACC TGGGGTGGAC CGCTTGCCCA TGGTGGCTGG  1380
ACCCCTATCC CCCCAACTGC TGACTTCCCC ATTCCCCAGT GTGGCATCCA GTGCCCCTCC  1440
CCTGACTGGC AAGCGAGGCC GGGGCCGCCC AAGGAAGGCC AACCTGCTGG ACTCAATGTT  1500
TGGGTCCCCA GGGGGCCTGA GGGAGGCAGG CATCCTTCCA TGCGGTCTAT GTGGTAAGGT  1560
GTTCACTGAT GCCAACCGGC TCCGGCAGCA CGAGGCCCAG CACGGTGTCA CCAGCCTCCA  1620
GCTGGGCTAC ATCGACCTTC CTCCTCCGAG GCTGGGTGAG AATGGGCTAC CCATCTCTGA  1680
AGACCCCGAC GGCCCCCGAA AGAGGAGCCG GACCAGGAAG CAGGTGGCTT GTGAGATCTG  1740
CGGCAAGATC TTCCGTGATG TGTATCATCT TAACCGGCAC AAGCTGTCCC ACTCTGGGGA  1800
GAAGCCCTAC TCCTGCCCTG TGTGTGGGTT GCGGTTCAAG AGAAAAGACC GCATGTCCTA  1860
CCATGTGCGG TCCCATGATG GGTCCGTGGG CAAGCCTTAC ATCTGCCAGA GCTGTGGGAA  1920
AGGCTTCTCC AGGCCTGATC ACTTGAACGG ACATATCAAG CAGGTGCACA CTTCTGAGCG  1980
GCCTCACAAG TGTCAGACCT GCAATGCTTC TTTTGCCACC CGAGACCGTC TGCGCTCCCA  2040
CCTGGCCTGT CATGAAGACA AGGTGCCCTG CCAGGTGTGT GGGAAGTACT TGCGGGCAGC  2100
ATACATGGCA GACCACCTGA AGAAGCACAG CGAGGGGCCC AGCAACTTCT GCAGTATCTG  2160
TAACCGAGGT TTCTCCTCTG CCTCCTACTT AAAGGTCCAT GTTAAAACCC ACCACGGTGT  2220
TCCCCTTCCC CAGGTCTCCA GGCACCAGGA GCCCATCCTG AATGGGGGAG CAGCGTTCCA  2280
CTGCGCCAGG ACCTATGGCA ACAAAGAAGG CCAGAAATGC TCACATCAGG ATCCGATTGA  2340
GAGCTCTGAC TCCTATGGTG ACCTCTCAGA TGCCAGCGAC CTGAAGACGC CAGAGAAGCA  2400
GAGTGCCAAT GGCTCTTTCT CCTGCGACAT GGCAGTCCCC AAAAACAAAA TGGAGTCTGA  2460
TGGGGAGAAG AAGTACCCAT GCCCTGAATG TGGGAGCTTC TTCCGCTCTA AGTCCTACTT  2520
GAACAAACAC ATCCAGAAGG TGCATGTCCG GGCTCTCGGG GGCCCCCTGG GGGACCTGGG  2580
CCCTGCCCTT GGCTCACCTT TCTCTCCTCA GCAGAACATG TCTCTCCTCG AGTCCTTTGG  2640
GTTTCAGATT GTTCAGTCGG CATTTGCGTC ATCTTTAGTA GATCCTGAGG TTGACCAGCA  2700
GCCCATGGGG CCTGAAGGGA AATGAGGCAG CTGCTGTGTC CCCACGGAAA CAACCATCTG  2760
GGGACTGCTG GGAAATGCTG TGAATGCGGA GGGAAGTGAT GTTTGGGTTC TGTAGCTGAG  2820
AGATTTTTAT TCATTTTTAA CTGCCCCCCA ACCCCACTCC AACTCCTTCT CCACCACCCA  2880
TTCTCCCAAT GGTCTTTAGA AATAGATTTT CATCTGATAT TCTGCAGAAA TATCAATGAG  2940
ACTTGGTATG GGACAGGGGC AGAAAACACT ACATAGGCCT CCAAGGCAAA ACCAGTCCCA  3000
GTTTCTTTAA TGGGAAGAAG CTGGAATTCC TGGTGCTCAA TTCTTAGTGA CCCCAATCCT  3060
ATACCCAAAT CTATGATATT CTGGGACCTC AGTGATTTTG GTCCCCTCCC ACTTCTCTAG  3120
TTCGTCATCC TCCCTTCCCA TATCCTTCAA AAGAACCACA CTAGGGTCTC CACCTACTTA  3180
TACAATGCGG ATGCCCAACT GTTTTTAAGG AAGCCAGAAG CATCCCATGG ACCATGGGGT  3240
GAGTGTCCTC CAAGAGCCCC CTGAGCTCAG CCCTCTGCCT GGAGGGGCTCC AGACCTTTCT  3300
GAGCCCTGCT TGGAGGCGAG CATTTTCACT GCTAGGACAA GCTCAGCTGT TGAGGACACC  3360
CCCACCCCAA ATTTCAGTTC TTACGTGATT TTAACCATTC AACATGCTGT TGGGTTTTAA  3420
TTCTCTAATT ATTATTATTA TTGTTATTAT TTTTTAGGAC CAGTTGTAGT GAATTGCTAC  3480
TGAAAGCTAT CCCAGGTGAT ACAGAGCTCT TTGTAAACCG CAGTCACACA TTAGGGTTAG  3540
TATTAAACTT TGTTTAGATG TACCATAATT AACTTGGCTA GTTGATTGTT TGAAGTCTAT  3600
GGAAGAAATA GTTTTATGCA AAATTTTAAA AAATGCCAGT CTGGTCAGGG AAGTAGGGGG  3660
TTTCAATGCT GTTGGGAACC AGGAAGGTGG GACAGCCGGC AGGTAGGGAC ATTGTGTACC  3720
TCAGTTGTGT CACATGTGAG CAAGCCCAGG TTGACCTTGT GATGTGAATT GATCTGATCA  3780
GACTGTATTA AAAATGTTAG TACATTACTC TA
```

```
MERVNDASCG PSGCYTYQVS RHSTEMLHNL NQQRKNGGRF CDVLLRVGDE SFPAHRAVLA   60
ACSEYFESVF SAQLGDGGAA DGGPADVGGA TAAPGGGAGG SRELEMHTIS SKVFGDILDF  120
AYTSRIVVRL ESFPELMTAA KFLLMRSVIE ICQEVIKQSN VQILVPPARA DIMLFRPPGT  180
SDLGFPLDMT NGAALAANSN GIAGSMQPEE EAARAAGAAI AGQASLPVLP GVDRLPMVAG  240
PLSPQLLTSP FPSVASSAPP LTGKRGRGRP RKANLLDSMF GSPGGLREAG ILPCGLCGKV  300
FTDANRLRQH EAQHGVTSLQ LGYIDLPPPR LGENGLPISE DPDGPRKRSR TRKQVACEIC  360
GKIFRDVYHL NRHKLSHSGE KPYSCPVCGL RFKRKDRMSY HVRSHDGSVG KPYICQSCGK  420
GFSRPDHLNG HIKQVHTSER PHKCQTCNAS FATRDRLRSH LACHEDKVPC QVCGKYLRAA  480
YMADHLKKHS EGPSNFCSIC NREGQKCSHQ DPIESSDSYG DLSDASDLKT PEKQSANGSF  540
SCDMAVPKNK MESDGEKKYP CPECGSFFRS KSYLNKHIQK VHVRALGGPL GDLGPALGSP  600
FSPQQNMSLL ESFGFQIVQS AFASSLVDPE VDQQPMGPEG K
```

```
1          11         21         31         41         51
|          |          |          |          |          |
CCCTACTCCG CCTCTCGGGA TCCTTTAAGA GGCGGGGCTT GGCTGCCAGC TCCGCGGCCC   60
GGGCAAAAGG CTGGGACTTT ACTCCGGGTG GCGGCGAGGA CGAGTCTGTG CTCCATCAGC  120
```

413

```
          TGCCGCACCC GCCGCCTCCC GCCCCCAAAC CCCATCCCCG CGGTTGAGCC ACGATGAGCG   180
          GCAGAGTCGG CGATCTGAGC CCCAGGCAGA AGGAGGCATT GGCCAAGTTT CGGGAGAATG   240
          TCCAGGATGT GCTGCCGGCC CTGCCGAATC CAGATGACTA TTTTCTCCTG CGTTGGCTCC   300
          GAGCCAGAAG CTTCGACCTG CAGAAGTCGG AGGCCATGCT CCGGAAGCAT GTGGAGTTCC   360
  5       GAAAGCAAAA GGACATTGAC AACATCATTA GCTGGCAGCC TCCAGAGGTG ATCCAACAGT   420
          ATCTGTCAGG GGGTATGTGT GGCTATGACC TGGATGGCTG CCCAGTCTGG TACGACATAA   480
          TTGGACCTCT GGATGCCAAG GGTCTGCTGT TCTCAGCCTC CAAACAGGAC CTGCTGAGGA   540
          CCAAGATGCG GGAGTGTGAG CTGCTTCTGC AAGAGTGTGC CCACCAGACC ACAAAGTTGG   600
          GGAGGAAGGT GGAGACCATC ACCATAATTT ATGACTGCGA GGGGCTTGGC CTCAAGCATC   660
  10      TCTGGAAGCC TGCTGTGGAG GCCTATGGAG AGTTTCTCTG CATGTTTGAG GAAAATTATC   720
          CCGAAACACT GAAGCGTCTT TTTGTTGTTA AAGCCCCCAA ACTGTTTCCT GTGGCCTATA   780
          ACCTCATCAA ACCCTTCCTG AGTGAGGACA CTCGTAAGAA GATCATGGTC CTGGGAGCAA   840
          ATTGGAAGGA GGTTTTACTG AAACATATCA GCCCTGACCA GGTGCCTGTG GAGTATGGGG   900
          GCACCATGAC TGACCCTGAT GGAAACCCCA AGTGCAAATC CAAGATCAAC TACGGGGGTG   960
  15      ACATCCCCAG GAAGTATTAT GTGCGAGACC AGGTGAAACA GCAGTATGAA CACAGCGTGC  1020
          AGATTTCCCG TGGCTCCTCC CACCAAGTGG AGTATGAGAT CCTCTTCCCT GGCTGTGTCC  1080
          TCAGGTGGCA GTTTATGTCA GATGGAGCGG ATGTTGGTTT TGGGATTTTC CTGAAGACCA  1140
          AGATGGGAGA GAGGCAGCGG GCAGGGGAGA TGACAGAGGT GCTGCCCAAC CAGAGGTACA  1200
          ACTCCCACCT GGTCCCTGAA GATGGGACCC TCACCTGCAG TGATCCTGGC ATCTATGTCC  1260
  20      TGCGGTTTGA CAACACCTAC AGCTTCATTC ATGCCAAGAA GGTCAATTTC ACTGTGGAGG  1320
          TCCTGCTTCC AGACAAAGCC TCAGAAGAGA AGATGAAACA GCTGGGGGCA GGCACCCCGA  1380
          AATAACACCT TCTCCTATAG CAGGCCTGGC CCCCTCAGTG TCTCCCTGTC AATTTCTACC  1440
          CCTTGTAGCA GTCATTTTCG CACAACCCTG AAGCCCAAAG AAACTGGGCT GGAGGACAGA  1500
          CCTCAGGAGC TTTCATTTCA GTTAGGCAGA GGAAGAGCGA CTGCAGTGGG TCTCCGTGTC  1560
  25      TATCAAATAC CTAAGGAGTC CCCAGGAGCT GGCTGGCCAT CGTGATAGGA TCTGTCTGTC  1620
          CTGTAAACTG TGCCAACTTC ACCTGTCCAG GGACAGCGAA GCTGGGGGTG GCGGGGGGCA  1680
          TGTACCACAG GGTGGCAGCA GGGAAAAAAA TTAGAAAAGG GTGAAAGATT GGGACTTAAC  1740
          ACTTCAGGGA AGTCAGCTGC CGGGGGAGAAA CTTGCTCCTA AATGAACACA TAAGTTTAGA  1800
          TCGCAATGAG GAGTAGCAGG GTAGCTGGTT GCTAGAGTTA CGGTGGGGAT CAGAAACTCT  1860
  30      TCCAAACATT TTAGCACTGA GGCTGGGGTA GCTTTTGGCT TTTCCCAGGT CTCAGGAGGT  1920
          GGCCTGAGTC AGCACACATC TTCCCACTCG GTAGACAGGC TGGCCTCTCC CTCACTTTGA  1980
          GACTTTGGCA ACTCCTGGGC CACACGGCCT GCCTCTTTGA TTACTAATGA TTGTCAGTGA  2040
          CTCAGAGCTT CCTGGGACTT CGGGTACCCA CCCGCTGTTC TCCATGCAAA CAAAGCGCCA  2100
          GGGAAATGAC CCACAGGGAT CGCAGCTGCA GGGAGGGCCA GGGAGGTTGG GGGTGGGAGT  2160
  35      GAATGCTAAA AGCAGATCGT CCAGTGCCCT TTTCAGTGCT ACCGGCCTCT CACCAAGCAG  2220
          TCCTCCATGT GAGCAACCCC GAGACAAAAA TGCTAAGTGG GATCAAGAGA GCAGCACTCG  2280
          GAGAGGGTGT TTGCCAGTCT GAGTGTCCCG CGGTGCCCGC CAACCCGCTT CCTGACTGAC  2340
          CTGAGCAAGG TCTTACTAAG CAGTCCCATC TCTGTGGGAG GCATGCAACG CGTGCAGGGA  2400
          GTTCAGGTGC CGGTCGGCGT AGCCAGGCCT GGAGGCCCCC CAGGCAGGAG GCCGCCCAAA  2460
  40      GGCGGGGCCG GCGTCTCGCA GACTAGGGGC TGGGGGCGGC CACAGACGGC CTCGAAACCA  2520
          CAGCCCTTAC CCCAATCCCA CGAGCCCCGC CAACGAACCA CAGGTGCTGG GCTTTAGAGA  2580
          ACATGGGAAG GCGGCCCCAG ACCTGGCGGG AACGCCTTTC CCTCAGAGCC AGGCCCCGGC  2640
          CCCGTCTGGG AAGCTCATCT TGCGAAGCTG AGGGAGCTCA GGGCAAAGGC CAGGCTAGCG  2700
          CGGACCGGAA GGGGCCGAGG CTGCACGGGC CTCTGCCAGA ACGCTCAGGA CATCCCGGCC  2760
  45      TGGGTTTACA ACGCTGTTAG GAAAATTAAC CAATGAATAA AGCAACGTTC AGTGCGCA
```

<u>SEQ ID NO:267 PBY9 Protein sequence:</u>
Protein Accession #:   NP_036561

```
          MSGRVGDLSP RQKEALAKFR ENVQDVLPAL PNPDDYFLLR WLRARSFDLQ KSEAMLRKHV   60
          EFRKQKDIDN IISWQPPEVI QQYLSGGMCG YDLDGCPVWY DIIGPLDAKG LLFSASKQDL  120
          LRTKMRECEL LLQECAHQTT KLGRKVETIT IIYDCEGLGL KHLWKPAVEA YGEFLCMFEE  180
          NYPETLKRLF VVKAPKLFPV AYNLIKPFLS EDTRKKIMVL GANWKEVLLK HISPDQVPVE  240
          YGGTMTDPDG NPKCKSKINY GGDIPRKYYV RDQVKQQYEH SVQISRGSSH QVEYEILFPG  300
          CVLRWQFMSD GADVGFGIFL KTKMGERQRA GEMTEVLPNQ RYNSHLVPED GTLTCSDPGI  360
          YVLRFDNTYS FIHAKKVNFT VEVLLPDKAS EEKMKQLGAG TPK
```

<u>SEQ ID NO:268 PBH8 DNA sequence</u>
Nucleic Acid Accession#:   XM_009756
Coding sequence: 301-1440 (underlined sequence corresponds to start and stop codon)

```
          1          11         21         31         41         51
          |          |          |          |          |          |
          GTGGGGACAG CCCGAGCCGCG CCGGGCCCCT GGACGGCGTC GCCAAGGAGC TGGGATCGCA   60
          CTTGCTGCAG ACTTTGGATG GATTTGTTTT TGTGGTAGCA TCTGATGGCA AAATCATGTA  120
          TATATCCGAG ACCGCTTCTG TCCATTTAGG CTTATCCCAG GTGGAGCTCA CGGGCAACAG  180
          TATTTATGAA TACATCCATC CTTCTGACCA CGATGAGATG ACCGCTGTCC TCACGGCCCA  240
          CCAGCCGCTG CACCACCACC TGCTCCAAGG TATGAGATAG AGAGGTCGTT CTTTCTTCGA  300
          ATGAAATGTG TCTTGGCGAA AAGGAACGCG GGCCTGACCT GCAGCGGATA CAAGGTCATC  360
          CACTGCAGTG GCTACTTGAA GATCAGGCAG TATATGCTGG ACATGTCCCT GTACGACTCC  420
          TGCTACCAGA TTGTGGGGCT GGTGGCCGTG GGCCAGTCGC TGCCACCCAG TGCCATCACC  480
          GAGATCAAGC TGTACAGTAA CATGTTCATG TTCAGGGCCA GCCTTGACCT GAAGCTGATA  540
          TTCCTGGATT CCAGGGTGAC CGAGGTGACG GGGTACGAGC CGCAGGACCT GATCGAGAAG  600
          ACCCTATACC ATCACGTGCA CGGCTGCGAC GTGTTCCACC TCCGCTACGC ACACCACCTC  660
          CTGTTGGTGA AGGGCCAGGT CACCACCAAG TACTACCGGC TGCTGTCCAA GCGGGGCGGC  720
          TGGGTGTGGG TGCAGAGCTA CGCCACCGTG GTGCACAACA GCCGCTCGTC CCGGCCCCAC  780
          TGCATCGTGA GTGTCAATTA TGTACTCACG GAGATTGAAT ACAAGGAACT TCAGCTGTCC  840
          CTGGAGCAGG TGTCCACTGC CAAGTCCCAG GACTCCTGGA GGACCGCCTT GTCTACCTCA  900
```
```
                                     414
```

```
CAAGAAACTA GGAAATTAGT GAAACCCAAA AATACCAAGA TGAAGACAAA GCTGAGAACA   960
AACCCTTACC CCCCACAGCA ATACAGCTCG TTCCAAATGG ACAAACTGGA ATGCGGCCAG   1020
CTCGGAAACT GGAGAGCCAG TCCCCCTGCA AGCGCTGCTG CTCCTCCAGA ACTGCAGCCC   1080
CACTCAGAAA GCAGTGACCT TCTGTACACG CCATCCTACA GCCTGCCCTT CTCCTACCAT   1140
TACGGACACT TCCCTCTGGA CTCTCACGTC TTCAGCAGCA AAAAGCCAAT GTTGCCGGCC   1200
AAGTTCGGGC AGCCCCAAGG ATCCCCTTGT GAGGTGGCAC GCTTTTTCCT GAGCACACTG   1260
CCAGCCAGCG GTGAATGCCA GTGGCATTAT GCCAACCCCC TAGTGCCTAG CAGCTCGTCT   1320
CCAGCTAAAA ATCCTCCAGA GCCACCGGCG AACACTGCTA GGCACAGCCT GGTGCCAAGC   1380
TACGAAGGCA AGCAGATGTC CTCTGCGGAG ATACCGCCAG CTCCCCAGGA CGCAGACTGA   1440
CTCCTGTTTG CTCGCTGGAC CAAC
```

SEQ ID NO:269 PBH8 Protein sequence:
Protein Accession #:   NP_005060

```
MKEKSKNAAK TRREKENGEF YELAKLLPLP SAITSQLDKA SIIRLTTSYL KMRAVFPEGL   60
GDAWGQPSRA GPLDGVAKEL GSHLLQTLDG FVFVVASDGK IMYISETASV HLGLSQVELT   120
GNSIYEYIHP SDHDEMTAVL TAHQPLHHHL LQEYEIERSF FLRMKCVLAK RNAGLTCSGY   180
KVIHCSGYLK IRQYMLDMSL YDSCYQIVGL VAVGQSLPPS AITEIKLYSN MFMFRASLDL   240
KLIFLDSRVT EVTGYEPQDL IEKTLYHHVH GCDVFHLRYA HHLLLVKGQV TTKYYRLLSK   300
RGGWVWVQSY ATVVHNSRSS RPHCIVSVNY VLTEIEYKEL QLSLEQVSTA KSQDSWRTAL   360
STSQETRKLV KPKNTKMKTK LRTNPYPPQQ YSSFQMDKLE CGQLGNWRAS PPASAAAPPE   420
LQPHSESSDL LYTPSYSLPF SYHYGHFPLD SHVFSSKKPM LPAKFGQPQG SPCEVARFFL   480
STLPASGECQ WHYANPLVPS SSSPAKNPPE PPANTARHSL VPSYEAPAAA VRRFGEDTAP   540
PSFPSCGHYR EEPALGPAKA ARQAARDGAR LALARAAPEC CAPPTPEAPG APAQLPFVLL   600
NYHRVLARRG PLGGAAPAAS GLACAPGGPE AATGALRLRH PSPAATSPPG APLPHYLGAS   660
VIITNGR
```

SEQ ID NO:270 PBJ9 DNA sequence:
Nucleic Acid Accession#:   AA760894

```
GGCACGAGGA GAAGATGTGG CTTGCTCATG CTTGACTTCT GCCATGGTTG TGAGGCCTCC   60
CCAGCCATGT GGAACTGTTT TCAGGTGCTG GTTCCATGGC TCTTCCTGAG CCGAAAATAA   120
GGAAACTCCA TAGACCTTGT CCACTGGAAC TCGTTCCCAT CTACCCTCCA CTCTATCCAG   180
GGTGATGGAT CTCTGCAGTA AGTGGAAGAG TTCTTCATGG CCCCCAAGGT TATATCCATC   240
TAGAACTTCA GCACGTAATT TCATCTGGAA ATAGTGCCTT TGTGGATATA AGTTAGGTAA   300
AACTGAAGAT GAGATCATAC TGGATTAGGA TGGGATCTAA ATCCAATGAA AATGTCTTCA   360
TAAAAAACAG GAAAGAACCC ATAGAAACAC AAGGAAGAAG GTCATGTGAA GATGGAGGCA   420
GAGATTGGAG GGATGCAGCC ACCGGCCCAG GAATGCCAGC AGCCACCCAG AAGCTGGAAG   480
GAAATGAGGG ATTCTCTCCT AGAACCTTTA GAGAGRACAT GGTCCTGTGA ACAGCTTGAT   540
TTTGGACTTG CCCATAGCTT GTATACTCTT ACTTTGGATA CAATTTTATC CAAACTTGGC   600
TAAACAGTTT CTCAGCCTAT GGAAAATTTA AAATGGAGAA GATTCAACTC GATTCTTACA   660
GATTCAAAGC AAGAAAATGA TGGGAACATA GGAGGAGACC AAGAAAGCCT ATAAAAAGCA   720
AAAATATGAA GTGAACATTG TGGTAGCTTT AAGATGTTTA GTGTAGCTGC AGGCACCCTA   780
TACACATGAA AACCCCCAAG GGGAATCCCC ATATCACAGT GTAGTGTGAT ATTTGACATT   840
YGTGATCATY TAGAGATGTA CAGAAAAGGT GAATCTGTGT TCTGTATATT CTGCCTAAGG   900
CAAAGAAATG TTTAGCTYTC TTTAAAATAG TTCCATAATT TTTTYTAAAA AGCTTTGCTT   960
GAAAACTGTA AGCTTCCCAT ATCTGGAGCA TTTCACTGTT TA AATATTTGGA TAAATATGTT   1020
ATCTTCTTAC TTGGACATTT CATGTGTTTA GGGATTGTYT TYTAAATTCT TCCTAATTCA   1080
TATAGCTGCT AACACTTCCC GCAGAGCTAA ACCATTACAG ANTATGAAAT AAAGACCCTA   1140
TTGATTTGAA CTTAAAAAAA AAAMAMAAA AAAAAAAAAA AAAAAAAAAT GA
```

SEQ ID NO:271 PBQ4 DNA sequence
Nucleic Acid Accession#:   AA149579
Coding sequence: 1-1363 (underlined sequence corresponds to start and stop codon)

```
1          11         21         31         41         51
|          |          |          |          |          |
ATGGAATCAA TCTCTATGAT GGGAAGCCCT AAGAGCCTTA GTGAAACTTG TTTACCTAAT   60
GGCATAAATG GTATCAAAGA TGCAAGGAAG GTCACTGTAG GTGTGATTGG AAGTGGAGAT   120
TTTGCCAAAT CCTTGACCAT TCGACTTATT AGATGCGGCT ATCATGTGGT CATAGGAAGT   180
AGAAATCCTA AGTTTGCTTC TGAATTTTTT CCTCATGTGG TAGATGTCAC TCATCATGAA   240
GATGCTCTCA CAAAAACAAA TATAATATTT GTTGCTATAC ACAGAGAACA TTATACCTCC   300
CTGTGGGACC TGAGACATCT GCTTGTGGGT AAAATCCTGA TTGATGTGAG CAATAACATG   360
AGGATAAACC AGTACCCAGA ATCCAATGCT GAATATTTGG CTTCATTATT CCCAGATTCT   420
TTGATTGTCA AAGGATTTAA TGTTGTCTCA GCTTGGGCAC TTCAGTTAGG ACCTAAGGAT   480
GCCAGCCGGC AGGTTTATAT ATGCAGCAAC AATATTCAAC CGCGACAACA GGTTATTGAA   540
CTTGCCCGCC AGTTGAATTT CATTCCCATT GACTTGGGAT CCTTATCATC AGCCAGAGAG   600
ATTGAAAATT TACCCCTACG ACTCTTTACT CTCTGGAGAG GGCCAGTGGT GGTAGCTATA   660
AGCTTGGCCA CATTTTTTTT CCTTTATTCC TTTGTCAGAG ATGTGATTCA TCCATATGCT   720
AGAAACCAAC AGAGTGACTT TTACAAAATT CCTATAGAGA TTGTGAATAA AACCTTACCT   780
ATAGTTGCCA TTACTTTGCT CTCCCTAGTA TACCTCGCAG GTCTTCTGGC AGCTGCTTAT   840
CAACTTTATT ACGGCACCAA GTATAGGAGA TTTCCACCTT GGTTGGAAAC CTGGTTACAG   900
TGTAGAAAAC AGCTTGGATT ACTAAGTTTT TTCTTCGCTA TGGTCCATGT TGCCTACAGC   960
CTCTGCTTAC CGATGAGAAG GTCAGAGAGA TATTTGTTTC TCAACATGGC TTATCAGCAG   1020
GTTCATGCAA ATATTGAAAA CTCTTGGAAT GAGGAAGAAG TTTGGAGAAT TGAAATGTAT   1080
ATCTCCTTTG GCATAATGAG CCTTGGCTTA CTTTCCCTCC TGGCAGTCAC TTCTATCCCT   1140
TCAGTGAGCA ATGCTTTAAA CTGGAGAGAA TTCAGTTTTA TTCAGTCTAC ACTTGGATAT   1200
```

<div align="center">415</div>

```
GTCGCTCTGC TCATAAGTAC TTTCCATGTT TTAATTTATG GATGGAAACG AGCTTTTGAG  1260
GAAGAGTACT ACAGATTTTA TACACCACCA AACTTTGTTC TTGCTCTTGT TTTGCCCTCA  1320
ATTGTAATTC TGGATCTTTT GCAGCTTTGC AGATACCCAG ACTGA
```

5 SEQ ID NO:272 PBQ4 Protein sequence:
Protein Accession #:   none

```
    1          11         21         31         41         51
    |          |          |          |          |          |
10  MESISMMGSP KSLSETCLPN GINGIKDARK VTVGVIGSGD FAKSLTIRLI RCGYHVVIGS   60
    RNPKFASEFF PHVVDVTHHE DALTKTNIIF VAIHREHYTS LWDLRHLLVG KILIDVSNNM  120
    RINQYPESNA EYLASLFPDS LIVKGFNVVS AWALQLGPKD ASRQVYICSN NIQARQQVIE  180
    LARQLNFIPI DLGSLSSARE IENLPLRLFT LWRGPVVVAI SLATFFFLYS FVRDVIHPYA  240
15  RNQQSDFYKI PIEIVNKTLP IVAITLLSLV YLAGLLAAAY QLYYGTKYRR FPPWLETWLQ  300
    CRKQLGLLSF FFAMVHVAYS LCLPMRRSER YLFLNMAYQQ VHANIENSWN EEEVWRIEMY  360
    ISFGIMSLGL LSLLAVTSIP SVSNALNWRE FSFIQSTLGY VALLISTFHV LIYGWKRAFE  420
    EEYYRFYTPP NFVLALVLPS IVILDLLQLC RYPD
```

20   SEQ ID NO:273 PBQ5 DNA SEQUENCE
Nucleic Acid Accession#:   NM_001973
Coding sequence: 150-1445 (underlined sequence corresponds to start and stop codon)

```
    1          11         21         31         41         51
    |          |          |          |          |          |
25  CCGCCGCCTT CTACTCCGCC GCGGGGGTCG CAGCGGCTGC CGCGCCGTCC TCGAGTTTCC   60
    AGCGTGAGGA GGAGGCTGAG GGCGGAGAGG CGCATCGTGT TCGAGGCGGA GACCGAGGGG  120
    GAGCCCCGCG CGCGGCGTCG CTCATTGCTA TGGACAGTGC TATCACCCTG TGGCAGTTCC  180
    TTCTTCAGCT CCTGCAGAAG CCTCAGAACA AGCACATGAT CTGTTGGACC TCTAATGATG  240
30  GGCAGTTTAA GCTTTTGCAG GCAGAAGAGG TGGCTCGTCT CTGGGGGATT CGCAAGAACA  300
    AGCCTAACAT GAATTATGAC AAACTCAGCC GAGCCCTCAG ATACTATTAT GTAAAGAATA  360
    TCATCAAAAA AGTGAATGGT CAGAAGTTTG TGTACAAGTT TGTCTCTTAT CCAGAGATTT  420
    TGAACATGGA TCCAATGACA GTGGGCAGGA TTGAGGGTGA CTGTGAAAGT TTAAACTTCA  480
    GTGAAGTCAG CAGCAGTTCC AAAGATGTGG AGAATGGAGG GAAAGATAAA CCACCTCAGC  540
35  CTGGTGCCAA GACCTCTAGC CGCAATGACT ACATACACTC TGGCTTATAT TCTTCATTTA  600
    CTCTCAACTC TTTGAACTCC TCCAATGTAA AGCTTTTCAA ATTGATAAAG ACTGAGAATC  660
    CAGCCGAGAA ACTGGCAGAG AAAAAAATCTC CTCAGGAGCC CACACCATCT GTCATCAAAT  720
    TTGTCACGAC ACCTTCCAAA AAGCCACCAG TTGAACCTGT TGCTGCCACC ATTTCAATTG  780
    GCCCAAGTAT TTCTCCATCT TCAGAAGAAA CTATCCAAGC TTTGGAGACA TTGGTTTCCC  840
40  CAAAACTGCC TTCCCTGGAA GCCCCAACCT CTGCCTCTAA CGTAATGACT GCTTTTGCCA  900
    CCACACCACC CATTTCGTCC ATACCCCCTT TGCAGGAACC TCCCAGAACA CCTTCACCAC  960
    CACTGAGTTC TCACCCAGAC ATCGACACAG ACATTGATTC AGTGGCTTCT CAGCCAATGG 1020
    AACTTCCAGA GAATTTGTCT CTGGAGCCTA AAGACCAGGA TTCAGTCTTG CTAGAAAAGG 1080
    ACAAAGTAAA TAATTCATCA AGATCCAAGA AACCCAAAGG GTTAGGACTG GCACCCACCC 1140
45  TTGTGATCAC GAGCAGTGAT CCAAGCCCAC TGGGAATACT GAGCCCATCT CTCCCTACAG 1200
    CTTCTCTTAC ACCAGCATTT TTTTCACAGA CACCCATCAT ACTGACTCCA AGCCCCTTGC 1260
    TCTCCAGTAT CCACTTCTGG AGTACTCTCA GTCCTGTTGC TCCCCTAAGT CCAGCCAGAC 1320
    TGCAAGGTGC TAACACACTT TTCCAGTTTC CTTCTGTACT GAACAGTCAT GGGCCATTCA 1380
    CTCTGTCTGG GCTGGATGGA CCTTCCACCC CTGGCCCATT TTCCCCAGAC CTACAGAAGA 1440
50  CATAACCTAT GCACTTGTGG AATGAGAGAA CCGAGGAACG AAGAAACAGA CATTCAACAT 1500
    GATTGCATTT GAAGTGAGCA ATTGATAGTT CTACAATGCT GATAATAGAC TATTGTGATT 1560
    TTTGCCATTC CCCATTGAAA ACATCTTTTT AGGATTCTCT TTGAATAGGA CTCAAGTTGG 1620
    ACTATATGTA TAAAAATGCC TTAATTGGAG TCTAAACTCC ACCTCCCTCT GTCTTTTCCT 1680
    TTTCTTTTTC TTTCCTTCCT TCCTTTTCTT TTCTCCTTTA AAAATATTTT GAGCTTTGTG 1740
55  CTGAAGAAGT TTTTGGTGGG CTTTAGTGAC TGTGCTTTGC AAAAGCAATT AAGAACAAAG 1800
    TTACTCCTTC TGGCTATTGG GACCCTTTGG CCAGGAAAAA TTATGCTTAG AATCTATTAT 1860
    TTAAAGAAGT ATTTGTGAAA TGAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 1920
    AAAAAAAAAA AAA
```

60

SEQ ID NO:274 PBQ5 Protein sequence:
Protein Accession #:   NP_001964

```
    MDSAITLWQF LLQLLQKPQN KHMICWTSND GQFKLLQAEE VARLWGIRKN KPNMNYDKLS  60
65  RALRYYYVKN IIKKVNGQKF VYKFVSYPEI LNMDPMTVGR IEGDCESLNF SEVSSSSKDV 120
    ENGGKDKPPQ PGAKTSSRND YIHSGLYSSF TLNSLNSSNV KLFKLIKTEN PAEKLAEKKS 180
    PQEPTPSVIK FVTTPSKKPP VEPVAATISI GPSISPSSEE TIQALETLVS PKLPSLEAPT 240
    SASNVMTAFA TTPPISSIPP LQEPPRTPSP PLSSHPDIDT DIDSVASQPM ELPENLSLEP 300
70  KDQDSVLLEK DKVNNSSRSK KPKGLGLAPT LVITSSDPSP LGILSPSLPT ASLTPAFFSQ 360
    TPIILTPSPL LSSIHFWSTL SPVAPLSPAR LQGANTLFQF PSVLNSHGPF TLSGLDGPST 420
    PGPFSPDLQK T
```

SEQ ID NO:275 PBY3 DNA SEQUENCE
75 Nucleic Acid Accession#:   AB040921
Coding sequence: 131-2560 (underlined sequence corresponds to start and stop codon)

```
    1          11         21         31         41         51
    |          |          |          |          |          |
    |          |          |          |          |          |
```

416

```
      AATCAGGAAC AGATCATATA TTGACCGAGA TTCTGAGTAT CTCTTGCAAG AAAATGAACC     60
      AGATGGAACT TTAGACCAAA AATTATTGGA AGATTTACAA AAGAAAAAAA ATGACCTTCG    120
      GTATATTGAA ATGCAGCATT TCAGAGAAAA GCTGCCTTCG TATGGAATGC AAAAGGAATT    180
      GGTAAATTTA ATTGATAACC ATCAGGTAAC AGTAATAAGT GGTGAAACTG GTTGTGGCAA    240
    AACCACTCAA GTTACTCAGT TCATTTTGGA TAACTACATT GAAAGAGGAA AAGGATCTGC    300
      TTGCAGAATA GTTTGTACTC AGCCAAGAAG AATTAGTGCC ATTTCAGTTG CGGAAAGAGT    360
      AGCTGCAGAA AGGGCAGAAT CTTGTGGCAG TGGTAATAGT ACTGGATATC AAATTCGTCT    420
      CCAGAGTCGG TTGCCAAGGA AACAGGGTTC TATCTTATAC TGTACAACAG GAATCATCCT    480
    TCAGTGGCTC CAGTCAGACC CGTATTTGTC CAGTGTTAGT CATATCGTAC TTGATGAAAT    540
      CCATGAAAGA AATCTGCAGT CAGATGTTTT AATGACTGTT GTTAAAGACC TTCTCAATTT    600
      TCGATCTGAC TTGAAAGTAA TATTGATGAG TGCAACATTG AATGCAGAAA AGTTTTCAGA    660
      ATATTTTGGT AACTGTCCAA TGATACATAT ACCTGGTTTT ACCTTTCCGG TTGTGGAATA    720
    TCTTTTGGAA GATGTAATTG AAAAAATAAG GTATGTTCCA GAACAAAAAG AACACAGATC    780
      CCAGTTTAAG AGGGGGTTTCA TGCAAGGGCA TGTAAATAGA CAAGAAAAAG AAGAAAAAGA    840
      AGCAATATAT AAAGAACGTT GGCCAGATTA TGTAAGGGAA CTGCGAAGAA GGTATTCTGC    900
      AAGTACTGTA GATGTTATAG AAATGATGGA GGATGATAAA GTTGATCTGA ATTTGATTGT    960
    TGCCCTCATC CGATACATTG TTTTGGAAGA AGAGGATGGT GCCGATACTGG TCTTTCTGCC   1020
      AGGCTGGGAC AATATCAGCA CTTTACATGA TCTCTTGATG TCACAAGTAA TGTTTAAATC   1080
      AGATAAATTT TTAATTATAC CTTTACATTC ACTGATGCCT ACAGTTAACC AGACACAGGT   1140
      GTTTAAAAGA ACCCCTCCTG GTGTTCGGAA AATAGTAATT GCTACCAACA TTGCGGAGAC   1200
    TAGCATTACC ATAGATGATG TCGTTTATGT GATAGATGGA GGAAAAATAA AAGAGACGCA   1260
      TTTTGATACT CAGAACAATA TCAGTACAAT GTCCGCTGAG TGGGTTAGTA AAGCTAATGC   1320
      CAAACAGAGA AAAGGTCGAG CTGGAAGAGT TCAACCTGGT CATTGCTATC ATCTGTATAA   1380
      TGGTCTTAGA GCAAGTCTTC TAGATGACTA TCAACTGCCA GAAATTTTGA GAACTCCTTT   1440
    GGAAGAACTT TGTTTACAAA TAAAGATTTT AAGGCTAGGT GGAATTGCTT ATTTTCTGAG   1500
      TAGATTAATG GACCCACCAT CAAATGAGGC AGTGTTACTC TCCATAAGAC ACCTGATGGA   1560
      GCTGAACGCT TTGGATAAAC AAGAAGAATT GACACCTCTT GGAGTCCACT TGGCACGATT   1620
      ACCCGTTGAG CCACATATTG GAAAAATGAT TCTTTTTGGA GCACTGTTCT GCTGCTTAGA   1680
    CCCAGTACTC ACTATTGCTG CTAGTCTCAG TTTCAAAGAT CCATTTGTCA TTCCACTGGG   1740
      AAAAGAAAAG ATTGCAGATG CAAGAAGAAA GGAATTGGCA AAGGATACTA GAAGTGATCA   1800
      CTTAACAGTT GTGAATGCGT TTGAGGGCTG GGAAGAGGCT AGGCGACGTG GTTTCAGATA   1860
      CGAAAAGGAC TATTGCTGGG AATATTTTCT GTCTTCAAAC ACACTGCAGA TGCTGCATAA   1920
    CATGAAAGGA CAGTTTGCTG AGCATCTTCT TGGAGCTGGA TTTGTAAGCA GTAGAAATCC   1980
      TAAAGATCCA GAATCTAATA TAAATTCAGA TAATGAGAAG ATAATTAAAG CTGTCATCTG   2040
      TGCTGGTTTA TATCCCAAAG TTGCTAAAAT TCGACTAAAT TTGGGTAAAA AAGAAAAAT   2100
      GGTAAAAGTT TACACAAAAA CCGATGGCCT GGTTGCTGTT CATCCTAAAT CTGTTAATGT   2160
    GGAGCAAACA GACTTTCACT ACAACTGGCT TATCTATCAC CTAAAGATGA GAACAAGCAG   2220
      TATATACTTG TATGACTGCA CAGAGGTTTC CCCATACTGT CTCTTGTTTT TTGGAGGTGA   2280
      CATTTCCATC CAGAAGGATA ACGATCAGGA AACTATTGCT GTAGATGAGT GGATTGTATT   2340
      TCAGTCTCCA GCAAGAATTG CCCATCTTGT TAAGGAATTA AGAAAGGAAC TAGATATTCT   2400
    TCTGCAAGAG AAGATTGAAA GTCCTCATCC TGTAGACTGG AATGACACTA AATCCAGAGA   2460
      CTGTGCAGTA CTGTCAGCTA TTATAGACTT GATCAAAACA CAGGAAAAGG CAACTCCCAG   2520
      GAACTTTCCG CCACGATTCC AGGATGGATA TTACAGCTGA CAGCTTTTCA GGGGTGGTCT   2580
      GAAAAGCCAG TTTGACAGCC ATTCTTCATC ATTGTTTAAA TTTTGGCTGG ATGCCAAACC   2640
    CTGGGACATG AACAATTTTC ATGTGTAAGG TAGAAGCCTT CAGTAGGTAG TAAAGACTTA   2700
      ATGTGCATGA CTTGATGTTA TATGTAGAGA TATATATATA TATATATATA CCATAAAAGC   2760
      AATATGTTCT CTGATCATAT ACTCTGCTGT GGTCATGCCC ACTCTTTGGG AGTATATTCC   2820
      CTTTATATAT ATTGAGTATT GTACCACTTG AGAAATTCCT TTGTTCTGTT ATACAAAATT   2880
    AATCTTTCTG CTCATAATGA TTGATGATAC CACCAGTAAA AATAGGATGT TTACCCCAAA   2940
      ACAAGTGTCA ATTAAGAATT TGAACACAAC CACATTTTTT AAAATGAAAC TTCTATCGGA   3000
      AGTAAATTAA TTTGTTGTAA TAAAGTCCAG TATTTAATAA AATGTACAAT GTTAAATCTC
```

SEQ ID NO:276 PBY3 Protein sequence:
Protein Accession #:   BAA96012

```
IRNRSYIDRD SEYLLQENEP DGTLDQKLLE DLQKKKNDLR YIEMQHFREK LPSYGMQKEL    60
VNLIDNHQVT VISGETGCGK TTQVTQFILD NYIERGKGSA CRIVCTQPRR ISAISVAERV   120
AAERAESCGS GNSTGYQIRL QSRLPRKQGS ILYCTTGIIL QWLQSDPYLS SVSHIVLDEI   180
HERNLQSDVL MTVVKDLLNF RSDLKVILMS ATLNAEKFSE YFGNCPMIHI PGFTFPVVEY   240
LLEDVIEKIR YVPEQKEHRS QFKRGFMQGH VNRQEKEEKE AIYKERWPDY VRELRRRYSA   300
STVDVGWIEMME DDKVDLNLIV ALIRYIVLEE EDGAILVFLP GWDNISTLHD LLMSQVMFKS   360
DKFLIIPLHS LMPTVNQTQV FKRTPPGVRK IVIATNIAET SITIDDVVYV IDGGKIKETH   420
FDTQNNISTM SAEWVSKANA KQRKGRAGRV QPGHCYHLYN GLRASLLDDY QLPEILRTPL   480
EELCLQIKIL RLGGIAYFLS RLMDPPSNEA VLLSIRHLME LNALDKQEEL TPLGVHLARL   540
PVEPHIGKMI LFGALFCCLD PVLTIAASLS FKDPFVIPLG KEKIADARRK ELAKDTRSDH   600
LTVVNAFEGW EEARRRGFRY EKDYCWEYFL SSNTLQMLHN MKGQFAEHLL GAGFVSSRNP   660
KDPESNINSD NEKIIKAVIC AGLYPKVAKI RLNLGKKRKM VKVYTKTDGL VAVHPKSVNV   720
EQTDFHYNWL IYHLKMRTSS IYLYDCTEVS PYCLLFFGGD ISIQKDNDQE TIAVDEWIVF   780
QSPARIAHLV KELRKELDIL LQEKIESPHP VDWNDTKSRD CAVLSAIIDL IKTQEKATPR   840
NFPPRFQDGY YS
```

SEQ ID NO:277 PBY6 DNA SEQUENCE

Nucleic Acid Accession#:   AA464018
Coding sequence: 64-1669(underlined sequence corresponds to start and stop codon)

```
GATTTTATCC TGGAACATTA CAGTGAAGAT GGCTATTTAT ATGAAGATGA AATTGCAGAT    60
CTTATGGATC TGAGACAAGC TTGTCGGACG CCTAGCCGGG ATGAGGCCGG GGTGGAACTG   120
```

417

```
CTGATGACAT ACTTCATCCA GCTGGGCTTT GTCGAGAGTC GATTCTTCCC GCCCACACGG  180
CAGATGGGAC TCCTGTTCAC CTGGTATGAC TCTCTCACCG GGGTTCCGGT CAGCCAGCAG  240
AACCTGCTGC TGGAGAAGGC CAGTGTCCTG TTCAACACTG GGGCCCTCTA CACCCAGATT  300
GGGACCCGGT GTGATCGGCA GACGCAGGCT GGGCTGGAGA GTGCCATAGA TGCCTTTCAG  360
AGAGCCGCAG GGGTTTTAAA TTACCTGAAA GACACATTTA CCCATACTCC AAGTTACGAC  420
ATGAGCCCTG CCATGCTCAG CGTGCTCGTC AAAATGATGC TTGCACAAGC CCAAGAAAGC  480
GTGTTTGAGA AAATCAGCCT TCCTGGGATC CGGAATGAAT TCTTCATGCT GGTGAAGGTG  540
GCTCAGGAGG CTGCTAAGGT GGGAGAGGTC TACCAACAGC TACACGCAGC CATGAGCCAG  600
GCGCCGGTGA AAGAGAACAT CCCCTACTCC TGGGCCAGCT TAGCCTGCGT GAAGGCCCAC  660
CACTACGCGG CCCTGGCCCA CTACTTCACT GCCATCCTCC TCATCGACCA CCAGGTGAAG  720
CCAGGCACGG ATCTGGACCA CCAGGAGAAG TGCCTGTCCC AGCTCTACGA CCACATGCCA  780
GAGGGGCTGA CACCCTTGGC CACACTGAAG AATGATCAGC AGCGCCGACA GCTGGGGAAG  840
TCCCACTTGC GCAGAGCCAT GGCTCATCAC GAGGAGTCGG TGCGGGAGGC CAGCCTCTGC  900
AAGAAGCTGC GGAGCATTGA GGTGCTACAG AAGGTGCTGT GTGCCGCACA GGAACGCTCC  960
CGGCTCACGT ACGCCCAGCA CCAGGAGGAG GATGACCTGC TGAACCTGAT CGACGCCCCC 1020
AGTGTTGTTG CTAAAACTGA GCAAGAGGTT GACATTATAT GCCCCAGTT CTCCAAGCTG 1080
ACAGTCACGG ACTTCTTCCA GAAGCTGGGC CCCTTATCTG TGTTTTCGGC TAACAAGCGG 1140
TGGACGCCTC CTCGAAGCAT CCGCTTCACT GCAGAAGAAG GGGACTTGGG GTTCACCTTG 1200
AGAGGGAACG CCCCCGTTCA GGTTCACTTC CTGGATCCTT ACTGCTCTGC CTCGGTGGCA 1260
GGAGCCCGGG AAGGAGATTA TATTGTCTCC ATTCAGCTTG TGGATTGTAA GTGGCTGACG 1320
CTGAGTGAGG TTATGAAGCT GCTGAAGAGC TTTGGCGAGG ACGAGATCGA GATGAAAGTC 1380
GTGAGCCTCC TGGACTCCAC ATCATCCATG CATAATAAGA GTGCCACATCA CTCCGTGGGA 1440
ATGCAGAAAA CGTACTCCAT GATCTGCTTA GCCATTGATG ATGACGACAA AACTGATAAA 1500
ACCAAGAAAA TCTCCAAGAA GCTTTCCTTC CTGAGTTGGG GCACCAACAA GAACAGACAG 1560
AAGTCAGCCA GCACCTTGTG CCTCCCATCG GTCGGGGCTG CACGGCCTCA GGTCAAGAAG 1620
AAGCTGCCCT CCCCTTTCAG CCTTCTCAAC TCAGACAGTT CTTGGTACTA A
```

SEQ ID NO:278 PBY6 Protein sequence:
Protein Accession #:   NP_149094

```
DFILEHYSED GYLYEDEIAD LMDLRQACRT PSRDEAGVEL LMTYFIQLGF VESRFFPPTR  60
QMGLLFTWYD SLTGVPVSQQ NLLLEKASVL FNTGALYTQI GTRCDRQTQA GLESAIDAFQ 120
RAAGVLNYLK DTFTHTPSYD MSPAMLSVLV KMMLAQAQES VFEKISLPGI RNEFFMLVKV 180
AQEAAKVGEV YQQLHAAMSQ APVKENIPYS WASLACVKAH HYAALAHYFT AILLIDHQVK 240
PGTDLDHQEK CLSQLYDHMP EGLTPLATLK NDQQRRQLGK SHLRRAMAHH EESVREASLC 300
KKLRSIEVLQ KVLCAAQERS RLTYAQHQEE DDLLNLIDAP SVVAKTEQEV DIILPQFSKL 360
TVTDFFQKLG PLSVFSANKR WTPPRSIRFT AEEGDLGFTL RGNAPVQVHF LDPYCSASVA 420
GAREGDYIVS IQLVDCKWLT LSEVMKLLKS FGEDEIEMKV VSLLDSTSSM HNKSATYSVG 480
MQKTYSMICL AIDDDDKTDK TKKISKKLSF LSWGTNKNRQ KSASTLCLPS VGAARPQVKK 540
KLPSPFSLLN SDSSWY
```


SEQ ID NO:279 PBY8 DNA SEQUENCE
Nucleic Acid Accession#:   AF107493
Coding sequence: 125-556 (underlined sequence corresponds to start and stop codon)

```
 1          11         21         31         41         51
 |          |          |          |          |          |
GAATTCGGCA CGAGCCTTGT TGGAGGTTCT GGGGCGCAGA ACCGCTACTG CTGCTTCGGT   60
CTCTCCTTGG GAAAAAATAA AATTTGAACC TTTTGGAGCT GTGTGCTAAA TCTTCAGTGG  120
GACAATGGGT TCAGACAAAA GAGTGAGTAG AACAGAGCGT AGTGGAAGAT ACGGTTCCAT  180
CATAGACAGG GATGACCGTG ATGAGCGTGA ATCCCGAAGC AGGCGGAGGG ACTCAGATTA  240
CAAAAGATCT AGTGATGATC GGAGGGGTGA TAGATATGAT GACTACCGAG ACTATGACAG  300
TCCGAGAGA GAGCGTGAAA GAAGGAACAG TGACCGATCC GAAGATGGCT ACCATTCAGA  360
TGGTGACTAT GGTGAGCACG ACTATAGGCA TGACATCAGT GACGAGAGGG AGAGCAAGAC  420
CATCATGCTG CGCGGCCTTC CCATCACCAT CACAGAGAGC GATATTCGAG AAATGATGGA  480
GTCCTTCGAA GGCCCTCAGC CTGCCGGATGT GAGGCTGATG AAGAGGAAAA CAGGTGAGAG  540
CTTGCTTAGT TCCTGATATT ATTGTTCTCT TCCCCATTCC CACCTCAGTC CCTAAAGAAC  600
ATCCTGATTC CCCCAGTCTT CAAGCACATG AATTCAGAAT GAAAGGTTTG CCATGGCTAA  660
GGAATGTGAC TCTTTGAAAA CCATGTTAGC ATCTGAGGAA CTTTTTTAAA CTTTGTTTTA  720
GGGACTTTTT TTTCCTTAGG TAAGTAATGA TTTATAAACT CCTTTTTTTT TTTGACTATA  780
GTCGGTTGCA TGGTTACTTT AAGCGTGGAA TCAAATGGAG TGGCATTTAG TTCAGGCGGC  840
TTGTTCCTTG CCATGGCAAA GTATCAAGAA GATCCCCAAG TCAAGTCACA TTTGTAAAGC  900
TGCTTCCCAA TTGGCTTTGT CACGCAGTGT TGAAGCAGTG GGAGGAGAGAT TCACCTGTTA  960
TAAAGGAACT GACTAACACA AGTATCCCGT CTATATCTGA ATGCTGTCTC TAGGTGTAAG 1020
CCGTGGTTTC GCCTTCGTGG AGTTTTATCA CTTGCAAGAT GCTACCAGCT GGATGGAAGC 1080
CAATCAGGTT GCTTCACTCA CCAAGTCTAG ATATTCATGA AAATGGAACA AGTCTGTACA 1140
ATTTTAAAAA AAGGTTGAAG GAGTGGTTTG TTCCAAAGGA GTGACTTTTT TTTAAAAAAA 1200
AAGCTTTGTA TATATTAAAA TTGATGTTAC TAGAATAAGT ACAGTACCAA GGACTTCATT 1260
ATAGAATTTG TTCTGCCTTT AAACATGGCT ACCTACCTGG CAGGGCTTTG TTAACTACTG 1320
AATACCTGTC TGGTAATCAC TAAAACATCT TTATGTTTCC CTTTTTTCTA GTTTGTTATA 1380
TTCCTATTAT GTCCATTGAG AGTAAGCTTA GTATATCAAA CTCTCCATTT GACAGTGAAG 1440
AGAACATAGT GAAAGTCTGT GGCGGCATTT TTATAAGTAA TTCCTTATTT CTGCCTGAAG 1500
ACCACAAAGC CTCCTGGAGG CGTAACTGCT CAGACCGGTC TTCAGGGAAT ATTTAAGGAC 1560
TTAGTGGAAT TTATGAACAA TAAGTCTGAT GAGATTAGCC TGGGAGTGGT GTCCTGCAGC 1620
TGTCTAATCT AGAGTGGCAT TAACATTCTA ATCTCCTTGA GAATGCCTTT TATAGTCTGT 1680
TCAAAGCAAG TCATTGATGG TTCTTCGAGG TAGTGTTAAC TGAAGTGTTC TTCAGTTTGT 1740
CAAGATAATG TTCAGTGCTT GGCACTTAAA TAACATTTTT TGCAAGAACT CCAAGGCACA 1800
```

418

```
TTATTGAATG CCTTTAACCA AGTGCATTCT GGGAAGTTTG CTTGACTCAT TATCTTGCTT   1860
TTCTGCAGCA TTCTGTGATT TGAGTCATCC ATGAATCCAT GAATAAAAGT TACATTCTTT   1920
GATTGGTAAT ATTGCCATTT ATAACAAGAC TCACTAATGA GGGTATCACT TTGACTGACT   1980
GATTTGTTAA AGTTTTTAAG CCTCTCATTT TCCTAACCCA GAAATCACAG CCTGATTTTA   2040
TTAAAAGTAG AGCTTCATTC ATTTCATACC ATAGATACCA TCCTAGTAAA TCCAGAACAT   2100
ATACAAGGTT CATGTGAGTC TGCTTTCTTG ACATGATAGC ATTGTTTGAT GCAGTGGATA   2160
TGTCAGAATG ACTAACCTAG GAGTTTGAAA CTCCTAAGAA ACTAAAACCT GTAAGACATT   2220
TAAAAGTCTC CACAATTTTA ATGTATACAA AGCTATGTTA CTGTGTAACA CATTACAGTT   2280
CAAATTCACT CCAGAAATAA AAGGCCAGTA GGATTAGGGA CTCACTGGTA GTTTGGAGTC   2340
TCCCAGCACA CATCCCTCCT AGTGGGATGA TCTATTCACA TATCTCCCAG CTTTTTTATT   2400
TTTGCTTCTG TATATCACAG TGAGTGGATG GCCCTTCAGC TTTTTCTCTC CTGGCCAGAC   2460
ATGCAGTCTT GCCTTTAGAT ATCGCAGAGA CAAAATTCAC AGCATGTCTT AAATCTTCCA   2520
GGATTTGCAA GAACCAAATT GCTCAACAGT ATGTATGTTT AGAGGGGTTA GACTCCTTTT   2580
TAAAATCTGG ATATCTAACC ACCTACTTAA ATCTGTTTGA TAGTGTCAAA CCACCCCCAC   2640
CCTTGATCCT CCCACCCCCA AAAAAAAAA AAAA
```

MGSDKRVSRT ERSGRYGSII DRDDRDERES RSRRRDSDYK RSSDDRRGDR YDDYRDYDSP 60
ERERERRNSD RSEDGYHSDG DYGEHDYRHD ISDERESKTI MLRGLPITIT ESDIREMMES 120
FEGPQPADVR LMKRKTGESL LSS

```
1         11        21        31        41        51
|         |         |         |         |         |
CGGCCGCTTT TTTCTCAAGA TGGCAGATTC CCACTGAGGC TGAGGGGGCC GAGCTCGCGC   60
GCCGCGTTCC CTTCTCCGTT GCCATGAACC GCGGACACCC CGGCCCCGAT GGCCCCCGTG   120
TACGAAGGTA TGGCCTCACA TGTGCAAGTT TTCTCCCCTC ACACCCTTCA ATCAAGTGCC   180
TTCTGTAGTG TGAAGAAACT AAAAGTAGAG CCAAGTTCCA ACTGGGACAT GACTGGGTAC   240
GGCTCCCACA GCAAAGTGTA CAGCCAGAGC AAGAACATAC CACCTTCTCA GCCAGCCTCC   300
ACAACCGTCA GCACCTCCTT GCCGGTCCCA AACCCAAGCC TACCTTACGA GCAGACCATC   360
GTCTTCCCAG GAAGCACCGG GCACATCGTG GTCACCTCAG CAAGCAGCAC TTCTGTCACC   420
GGGCAAGTCC TCGGCGGACC ACACAACCTA ATGCGTCGAA GCACTGTGAG CCTCCTTGAT   480
ACCTACCAAA AATGTGGACT CAAGCGTAAG AGCGAGGAGA TCGAGAACAC AAGCAGCGTG   540
CAGATCATCG AGGAGCATCC ACCCATGATT CAGAATAATG CAAGCGGGGC CACTGTCGCC   600
ACTGCCACCA CGTCTACTGC CACCTCCAAA AACAGCGGCT CCAACAGCGA GGGCGACTAT   660
CAGCTGGTGC AGCATGAGGT GCTGTGCTCC ATGACCAACA CCTACGAGGT CTTAGAGTTC   720
TTGGGCCGAG GGACGTTTGG ACAAGTGGTC AAGTGCTGGA AACGGGGCAC CAATGAGATC   780
GTAGCCATCA AGATCCTGAA GAACCGCCCA TCCTATGCCC GACAAGGTCA GATTGAAGTG   840
AGCATCCTGG CCCGGTTGAG CACGGAGAGT GCCGATGACT ATAACTTCGT CCGGGCCTAC   900
GAATGCTTCC AGCACAAGAA CCACACGTGC TTGGTCTTCG AGATGTTGGA GCAGAACCTC   960
TATGACTTTC TGAAGCAAAA CAAGTTTAGC CCCTTGCCCC TCAAATACAT TCGCCCCAGTT   1020
CTCCAGCAGG TAGCCACAGC CCTGATGAAA CTCAAAAGCC TAGGTCTTAT CCACGCTGAC   1080
CTCAAACCAG AAAACATCAT GCTGGTGGAT CCATCTAGAC AACCATACAG AGTCAAGGTC   1140
ATCGACTTTG GTTCAGCCAG CCACGTCTCC AAGGCTGTGT GCTCCACCTA CTTGCAGTCC   1200
AGATATTACA GGGCCCCTGA GATCATCCTT GGTTTACCAT TTTGTGAGGC AATTGACATG   1260
TGGTCCCTGG GCTGTGTTAT TGCAGAATTG TTCCTGGGTT GGCCGTTATA TCCAGGAGCT   1320
TCGGAGTATG ATCAGATTCG GTATATTTCA CAAACACAGG GTTTGCCTGC TGAATATTTA   1380
TTAAGCGCCG GGACAAAGAC AACTAGGTTT TTCAACCGTG ACACGGACTC ACCATATCCT   1440
TTGTGGAGAC TGAAGACACC AGATGACCAT GAAGCAGAGA CAGGGATTAA GTCAAAAGAA   1500
GCAAGAAAGT ACATTTTCAA CTGTTTAGAT GATATGGCCC AGGTGAACAT GACGACAGAT   1560
TTGGAAGGGA GCGACATGTT GGTAGAAAAG GCTGACCGGC GGGAGTTCAT TGACCTGTTG   1620
AAGAAGATGC TGACCATTGA TGCTGACAAG AGAATCACTC CAATCGAAAC CCTGAACCAT   1680
CCCTTTGTCA CCATGACACA CTTACTCGAT TTTCCCCACA GCACACACGT CAAATCATGT   1740
TTCCAGAACA TGGAGATCTG CAAGCGTCGG GTGAATATGT ATGACACGGT GAACCAGAGC   1800
AAAACCCCTT TCATCACGCA CGTGGCCCCC AGCACGTCCA CCAACCTGAC CATGACCTTT   1860
AACAACCAGC TGACCACTGT CCACAACCAG GCTCCCTCCT CTACCAGTGC CACTATTTCC   1920
TTAGCCAATC CCGAAGTCTC CATACTAAAC TACCCATCTA CACTCTACCA GCCCTCAGCG   1980
GCATCCATGG CTGCAGTGGC CCAGCGGAGC ATGCCCCTGC AGACAGGAAC AGCCCAGATT   2040
TGTGCCCGGC CTGACCCGTT CCAGCAAGCT CTCATCGTGT GTCCCCCCGG CTTCCAAGGC   2100
TTGCAGGCCT CTCCCTCTAA GCACGCTGGC TACTCGGTGC GAATGGAAAA TGCAGTTCCC   2160
ATCGTCACTC AAGCCCCAGG AGCTCAGCCT CTTCAGATCC AACCAGGTCT GCTTGCCCAG   2220
CAGGCTTGGC CAAGTGGGAC CCAGCAGATC CTGCTTCCCC CAGCATGGCA GCAACTGGACT   2280
GGAGTGGCCA CCCACACATC AGTGCAGCAT GCCACCGTGA TTCCCGAGAC CATGGCAGGC   2340
ACCCAGCAGC TGGCGGACTG GAGAAATACG CATGCTCACG GAAGCCATTA TAATCCCATC   2400
ATGCAGCAGC CTGCACTATT GACCGGTCAT GTGACCCTTC CAGCAGCACA GCCCTTAAAT   2460
GTGGGTGTGG CCCACGTGAT GCGGCAGCAG CCAACCAGCA CCACCTCCTC CCGGAAGAGT   2520
AAGCAGCACC AGTCATCTGT GAGAAATGTC TCCACCTGTG AGGTGTCCTC CTCTCAGGCC   2580
ATCAGCTCCC CACAGCGATC CAAGCGTGTC AAGGAGAACA CACCTCCCCG CTGTGTCCATG   2640
GTGCACAGTA GCCCGGCCTG CAGCACCTCG GTCACCTGTG GGTGGGGCGA CGTGGCCTCC   2700
AGCACCACCC GGGAACGGCA GCGGCAGACA ATTGTCATTC CCGACACTCC CAGCCCCACG   2760
GTCAGCGTCA TCACCATCAG CAGTGACACG GACGAGGAGG AGGAACAGAA ACACGCCCCC   2820
ACCAGCACTG TCTCCAAGCA AAGAAAAAAC GTCATCAGCT GTGTCACAGT CCACGACTCC   2880
CCCTACTCCG ACTCCTCCAG CAACACCAGC CCCTACTCCG TGCAGCAGCG TGCTGGGCAC   2940
```

```
AACAATGCCA ATGCCTTTGA CACCAAGGGG AGCCTGGAGA ATCACTGCAC GGGGAACCCC    3000
CGAACCATCA TCGTGCCACC CCTGAAAACC CAGGCCAGCG AAGTATTGGT GGAGTGTGAT    3060
AGCCTGGTGC CAGTCAACAC CAGTCACCAC TCGTCCTCCT ACAAGTCCAA GTCCTCCAGC    3120
AACGTGACCT CCACCAGCGG TCACTCTTCA GGGAGCTCAT CTGGAGCCAT CACCTACCGG    3180
CAGCAGCGGC CGGGCCCCCA CTTCCAGCAG CAGCAGCCAC TCAATCTCAG CCAGGCTCAG    3240
CAGCACATCA CCACGGACCG CACTGGGAGC CACCGAAGGC AGCAGGCCTA CATCACTCCC    3300
ACCATGGCCC AGGCTCCGTA CTCCTTCCCG CACAACAGCC CCAGCCACGG CACTGTGCAC    3360
CCGCATCTGG CTGCAGCCGC TGCCGCTGCC CACCTCCCCA CCCAGCCCCA CCTCTACACC    3420
TACACTGCGC CGGCGGCCCT GGGCTCCACC GGCACCGTGG CCCACCTGGT GGCCTCGCAA    3480
GGCTCTGCGC GCCACACCGT GCAGCACACT GCCTACCCAG CCAGCATCGT CCACCAGGTC    3540
CCCGTGAGCA TGGGCCCCCG GGTCCTGCCC TCGCCCACCA TCCACCCGAG TCAGTATCCA    3600
GCCCAATTTG CCCACCAGAC CTACATCAGC GCCTCGCCAG CCTCCACCGT CTACACTGGA    3660
TACCCACTGA GCCCCGCCAA GGTCAACCAG TACCCTTACA TATAAACACT GGAGGGGAGG    3720
GAGGGAGGGA GGGAGGGAGA GAATGGCCCG AGGGAGGAGG GAGAGAAGGA GGGAGGCGCT    3780
CCTGGGACCG TGGGCGCTGG CCTTTTATAC TGAAGATGCC GCACACAAAC AATGCAAACG    3840
GGGCAGGGGC GGGGGGGGGG GGGGCAGAGG GCAGGGGGAC GGGTCGGGAC ACCAGTGAAA    3900
CTTGAACCGG GAAGTGGGAG GACGTAGAGC AGAGAAGAGA ACATTTTTAA AAGGAAGGGA    3960
TTAAAGAGGG TGGGAAATCT ATGGTTTTTA TTTTAAAAAA
```

SEQ ID NO:282 PCI2 Protein sequence:
Protein Accession #: NP_073577

```
MAPVYEGMAS HVQVFSPHTL QSSAFCSVKK LKVEPSSNWD MTGYGSHSKV YSQSKNIPPS    60
QPASTTVSTS LPVPNPSLPY EQTIVFPGST GHIVVTSASS TSVTGQVLGG PHNLMRRSTV   120
SLLDTYQKCG LKRKSEEIEN TSSVQIIEEH PPMIQNNASG ATVATATTST ATSKNSGSNS   180
EGDYQLVQHE VLCSMTNTYE VLEFLGRGTF GQVVKCWKRG TNEIVAIKIL KNRPSYARQG   240
QIEVSILARL STESADDYNF VRAYECFQHK NHTCLVFEML EQNLYDFLKQ NKFSPLPLKY   300
IRPVLQQVAT ALMKLKSLGL IHADLKPENI MLVDPSRQPY RVKVIDFGSA SHVSKAVCST   360
YLQSRYYRAP EIILGLPFCE AIDMWSLGCV IAELFLGWPL YPGASEYDQI RYISQTQGLP   420
AEYLLSAGTK TTRFFNRDTD SPYPLWRLKT PDDHEAETGI KSKEARKYIF NCLDDMAQVN   480
MTTDLEGSDM LVEKADRREF IDLLKKMLTI DADKRITPIE TLNHPFVTMT HLLDFPHSTH   540
VKSCFQNMEI CKRRVNMYDT VNQSKTPFIT HVAPSTSTNL TMTFNNQLTT VHNQAPSSTS   600
ATISLANPEV SILNYPSTLY QPSAASMAAV AQRSMPLQTG TAQICARPDP FQQALIVCPP   660
GFQGLQASPS KHAGYSVRME NAVPIVTQAP GAQPLQIQPG LLAQQAWPSG TQQILLPPAW   720
QQLTGVATHT SVQHATVIPE TMAGTQQLAD WRNTHAHGSH YNPIMQQPAL LTGHVTLPAA   780
QPLNVGVAHV MRQQPTSTTS SRKSKQHQSS VRNVSTCEVS SSQAISSPQR SKRVKENTPP   840
RCAMVHSSPA CSTSVTCGWG DVASSTTRER QRQTIVIPDT PSPTVSVITI SSDTDEEEEQ   900
KHAPTSTVSK QRKNVISCVT VHDSPYSDSS SNTSPYSVQQ RAGHNNANAF DTKGSLENHC   960
TGNPRTIIVP PLKTQASEVL VECDSLVPVN TSHHSSSYKS KSSSNVTSTS GHSSGSSSGA  1020
ITYRQQRPGP HFQQQQPLNL SQAQQHITTD RTGSHRRQQA YITPTMAQAP YSFPHNSPSH  1080
GTVHPHLAAA AAAAHLPTQP HLYTYTAPAA LGSTGTVAHL VASQGSARHT VQHTAYPASI  1140
VHQVPVSMGP RVLPSPTIHP SQYPAQFAHQ TYISASPAST VYTGYPLSPA KVNQYPYI
```

SEQ ID NO:283 PBY1 DNA SEQUENCE
Nucleic Acid Accession#: NM_017700
Coding sequence: 147-806 (underlined sequence corresponds to start and stop codon)

```
1          11         21         31         41         51
|          |          |          |          |          |
AGTCACAGCC AGGTAACCCT GGAGTGAAGC GGTTTAGTTA GAAGGGAGCA GATAAACTCG     60
TCACTCTAGT AGCTTTAACC CTCACCCTGG GGCACCTTAG CAATCAGCCA TTGCCTGCAA    120
GCCTCCAAAG CTTGTCTTTG CCTAATATGG AGCCCAAAGA AGCCACTGGG AAAGAAAACA    180
TGGTCACCAA GAAAAAGAAT CTGGCCTTCT TGAGGTCTAG ACTCTATATG CTGGAGAGAA    240
GGAAGACTGA CACTGTGGTT GAGAGCAGTG TTTCTGGGGA CCACTCTGGC ACCTTGAGGA    300
GGAGCCAATC TGACAGGACC GAATACAACC AGAAATTACA AGAAAAGATG ACTCCACAGG    360
GTGAGTGTTC TGTAGCTGAG ACCTTAACCC CAGAGGAAGA GCATCATATG AAGAGGATGA    420
TGGCAAAGCG GGAAAAGATC ATTAAGGAGC TGATACAGAC AGAAAAGGAT TATCTCAATG    480
ATCTAGAGCT GTGTGTTAGG GAAGTGGTTC AGCCCCTGAG AAATAAAAAG ACTGATAGGC    540
TGGATGTGGA TAGCTTGTTT AGCAACATTG AGTCCGTGCA TCAGATATCG GCCAAGCTGC    600
TGTCATTGTT GGAAGAGGCC ACAACAGACG TGGAACCGGC CATGCAAGTA ATTGGAGAAG    660
TATTCTTGCA GATTAAAGGG CCACTGGAAG ATATTTATAA AATCTACTGC TATCACCATG    720
ATGAAGCACA TAGTATACTG GAGTCCTATG AAAAGGAAGA AGAGCTGAAG GAACATTTGA    780
GCCACTGTAT CCAGTCCTTA AAGTAAGGCC TTTTCAAATG ATGATTCCCA TCTCCTCTCA    840
GTTGCCTAGC AGGGAACATT TTAAATGGAT GTAGATGAAA GGTCTCACAT AAATCCTATG    900
TTTTATGAGA CTTGCTGGGA GCTCTGCTTT GCATTCCCTT TATAAAAAGC TGACATGCCA    960
GAAGCCCTGA TTGACTTTTT TTCCCCCTGC GAGAATGACT AAAAATAACA TGGAAGAAGA   1020
TTTAGAGCTC TGCAGCGATT GAAAAATGCA ATATCAAAAT ATAAAATGTG GAAGAAAAGC   1080
CTCTTCTTAA AGCTATTGTA ACTTGCCTGG CCCCACGTAG TTCAAGGATT ATGTGAGATA   1140
ACACGTGGCC CCATGACCAC TGGAGCACAT GGGTTAATGG AGTTAGGGGA ATGGCCTACA   1200
ACTCTGCATG GCCGTCTTCT TTCCCCAAAC TCACTGTGGG GAGATGGGTG AAGACAAGTC   1260
AGGCCTTGTT AAAGTTAGTT TCAGAACAAT TACTCATGCC TTCCTTTCTC ATCCCTAAAA   1320
CATTGGTGGG GGAGCTACAC AATGTACTTT TTCTTTTCTA GAGGAAGTAT CTATTCACTG   1380
TGAAAATCTG AAAAATATAA CAAAGTATGT GTAAGATAAA AACCCCTTGC TATTTCAAAA   1440
AAAAAAAAAA AAAAAAAAAA AAAA
```

SEQ ID NO:284 PBY1 Protein sequence:
Protein Accession #:   NP_060170

```
1          11         21         31         41         51
```

420

```
        |          |          |          |          |          |
5   MEPKEATGKE NMVTKKKNLA FLRSRLYMLE RRKTDTVVES SVSGDHSGTL RRSQSDRTEY    60
    NQKLQEKMTP QGECSVAETL TPEEEHHMKR MMAKREKIIK ELIQTEKDYL NDLELCVREV   120
    VQPLRNKKTD RLDVDSLFSN IESVHQISAK LLSLLEEATT DVEPAMQVIG EVFLQIKGPL   180
    EDIYKIYCYH HDEAHSILES YEKEEELKEH LSHCIQSLK
```

SEQ ID NO:285 PBQ9 DNA SEQUENCE

Nucleic Acid Accession#:   X66534
10 Coding sequence: 523-2676 (underlined sequence corresponds to start and stop codon)

```
     1          11         21         31         41         51
     |          |          |          |          |          |
15  CCCTTATGGC GATTGGGCGG CTGCAGAGAC CAGGACTCAG TTCCCCTGCC CTAGTCTGAG    60
    CCTAGTGGGT GGGACTCAGC TCAGAGTCAG TTTTCAGAAG CAGGTTTCAG TTGCAGAGTT   120
    TTCCTACACT TTTCCTGCGC TAGAGCAGCG AGCAGCCTGG AACAGACCCA GGCGGAGGAC   180
    ACCTGTGGGG GAGGGAGCGC CTGGAGGAGC TTAGAGACCC CAGCCGGGCG TGATCTCACC   240
    ATGTGCGGAT TTGCGAGGCG CGCCCTGGAG CTGCTAGAGA TCCGGAAGCA CAGCCCCGAG   300
    GTGTGCGAAG CCACCAAGAC TGCGGCTCTT GGAGAAAGCG TGAGCAGGGG GCCACCGCGG   360
    TCTCCGGCCT GTCTGCACCC TGTCGCCTGA GCTGCCTGAC AGTGACAATG ACATCCCAGT   420
    TACCAGTGTC CTTGAATTGA TAGTGGCTTC TGTTTGTCAG TCTCATATAA GAACTACAGC   480
    TCATCAGGAG GAGATCGCAG CAGGGTAAGA GACACCAACA CCATGTTCTG CACGAAGCTC   540
    AAGGATCTCA AGATCACAGG AGAGTGTCCT TTCTCCTTAC TGGCACCAGG TCAAGTTCCT   600
    AACGAGTCTT CAGAGGAGGC AGCAGGAAGC TCAGAGAGCT GCAAAGCAAC CGTGCCCATC   660
    TGTCAAGACA TTCCTGAGAA GAACATACAA GAAAGTCTTC CTCAAAGAAA AACCAGTCGG   720
    AGCCGAGTCT ATCTTCACAC TTTGGCAGAG AGTATTTGCA AACTGATTTT CCCAGAGTTT   780
    GAACGGCTGA ATGTTGCACT TCAGAGAACA TTGGCAAAGC ACAAAATAAA AGAAAGCAGG   840
    AAATCTTTGG AAAGAGAAGA CTTTGAAAAA ACAATTGCAG AGCAAGCAGT GCAGCAGAGT   900
    CCAGTGGAGT TATCAAAGAA TCTCTTGGTG AAGAGGTTTT TAAAATATGT TACGAGGAAG   960
    ATGAAAACAT CCTTGGGGTG GTTGGAGGCA CCCTTAAAGA TTTTTAAACA GCTTCAGTAC  1020
    CCTTCTGAAA CAGAGCAGCC ATTGCCAAGA AGCAGGAAAA AGGGGCAGCT TGAGGACGCC  1080
    TCCATTCTAT GCCTGGATAA GGAGGATGAT TTTCTACATG TTTACTACTT CTTCCCTAAG  1140
    AGAACCACCT CCCTGATTCT TCCCGGCATC ATAAAGGCAG CTGCTCACGT ATTATATGAA  1200
    ACGGAAGTGG AAGTGTCGTT AATGCCTCCC TGCTTCCATA ATGATTGCAG CGAGTTTGTG  1260
    AATCAGCCCT ACTTGTTGTA CTCCGTTCAC ATGAAAAGCA CCAAGCCATC CCTGTCCCCC  1320
    AGCAAACCCC AGTCCTCGCT GGTGATTCCC ACATCGCTAT TCTGCAAGAC ATTTCCATTC  1380
    CATTTCATGT TTGACAAAGA TATGACAATT CTGCAATTTG GCAATGGCAT CAGAAGGCTG  1440
    ATGAACAGGA GAGACTTTCA AGGAAAGCCT AATTTTGAAT ACTTTGAAAT TCTGACTCCA  1500
    AAAATCAACC AGACCTTTAG CGGGATCATG ACTATGTTGA ATATGCAGTT TGTTGTACGA  1560
    GTGAGGAGAT GGGACAACTC TGTGAAGAAA TCTTCAAGGG TTATGGACCT CAAAGGCCAA  1620
    ATGATCTACA TTGTTGAATC CAGTGCAATC TTGTTTTTGG GGTCACCCTG TGTGGACAGA  1680
    TTAGAAGATT TTACAGGACG AGGGCTCTAC CTCTCAGACA TCCCAATTCA CAATGCACTG  1740
    AGGGATGTGG TCTTAATAGG GGAACAAGCC CGAGCTCAAG ATGGCCTGAA GAAGAGGCTG  1800
    GGGAAGCTGA AGGCTACCCT TGAGCAAGCC CACCAAGCCC TGGAGGAGGA GAAGAAAAAG  1860
    ACAGTAGACC TTCTGTGCTC CATATTTCCC TGTGAGGTTG CTCAGCAGCT GTGGCAAGGG  1920
    CAAGTTGTGC AAGCCAAGAA GTTCAGTAAT GTCACCATGC TCTTCTCAGA CATCGTTGGG  1980
    TTCACTGCCA TCTGCTCCCA GTGCTCACCG CTGCAGGTCA TCACCATGCT CAATGCACTG  2040
    TACACTCGCT TCGACCAGCA GTGTGGAGAG CTGGATGTCT ACAAGGTGGA GACCATTGCG  2100
    ATGCCTATTG TGTGGCTTGG GGGATTACAC AAAGAGAGTG ATACTCATGC TGTTCAGATA  2160
    GCGCTGATGG CCCTGAAGAT GATGGAGCTC TCTGATGAAG TTATGTCTCC CCATGGAGAA  2220
    CCTATCAAGA TGCGAATTGG ACTGCACTCT GGATCAGTTT TTGCTGGCGT CGTTGGAGTT  2280
    AAAATGCCCC GTTACTGTCT TTTTGGAAAC AATGTCACTC TGGCTAACAA ATTTGAGTCC  2340
    TGCAGTGTAC CACGAAAAAT CAATGTCAGC CCAACAACTT ACAGATTACT CAAAGACTGT  2400
    CCTGGTTTCG TGTTTACCCC TCGATCAAGG GAGGAACTTC CACCAAACTT CCCTAGTGAA  2460
    ATCCCCGGAA TCTGCCATTT TCTGGATGCT TACCAACAAG GAACAAACTC AAAACCATGC  2520
    TTCCAAAAGA AAGATGTGGA AGATGCAAGC CAATTTTTTA GGCAAAGCAT CAGGAATAGA  2580
    TTAGCAACCT ATATACCTAT TTATAAGTCT TTGGGGTTTG ACTCATTGAA GATGTGTAGA  2640
    GCCTCTGAAA GCACTTTAGG GATTGTAGAT GGCTAACAAG CAGTATTAAA ATTTCAGGAG  2700
    CCAAGTCACA ATCTTTCTCC TGTTTAACAT GACAAAATGT ACTCACTTCA GTACTTCAGC  2760
    TCTTCAAGAA AAAAAAAAAA ACCTTAAAAA GCTACTTTTG TGGGAGTATT TCTATTATAT  2820
    AACCAGCACT TACTACCTGT ACTCAAAATT CAGCACCTTG TACATATATC AGATAATTGT  2880
    AGTCAATTGT ACAAACTGAT GGAGTCACCT GCAATCTCAT ATCCTGGTGG AATGCCATGG  2940
    TTATTAAAGT GTGTTTGTGA TAGTTGTCGT CAAAAAAAAA AAAAAAAAA AAAAAAAAAA  3000
    AAAA
```

SEQ ID NO:286 PBQ9 Protein sequence:
Protein Accession #:   Q02108

```
     1          11         21         31         41         51
     |          |          |          |          |          |
70  MFCTKLKDLK ITGECPFSLL APGQVPNESS EEAAGSSESC KATVPICQDI PEKNIQESLP    60
    QRKTSRSRVY LHTLAESICK LIFPEFERLN VALQRTLAKH KIKESRKSLE REDFEKTIAE   120
    QAVAAGVPVE VIKESLGEEV FKICYEEDEN ILGVVGGTLK DFLNSFSTLL KQSSHCQEAG   180
75  KRGRLEDASI LCLDKEDDFL HVYYFFPKRT TSLILPGIIK AAAHVLYETE VEVSLMPPCF   240
    HNDCSEFVNQ PYLLYSVHMK STKPSLSPSK PQSSLVIPTS LFCKTFPFHF MFDKDMTILQ   300
    FGNGIRRLMN RRDFQGKPNF EEYFEILTPK INQTFSGIMT MLNMQFVVRV RRWDNSVKKS   360
    SRVMDLKGQM IYIVESSAIL FLGSPCVDRL EDFTGRGLYL SDIPIHNALR DVVLIGEQAR   420
    AQDGLKKRLG KLKATLEQAH QALEEEKKKT VDLLCSIFPC EVAQQLWQGQ VVQAKKFSNV   480
    TMLFSDIVGF TAICSQCSPL QVITMLNALY TRFDQQCGEL DVYKVETIGD AYCVAGGLHK   540
```

421

ESDTHAVQIA LMALKMMELS DEVMSPHGEP IKMRIGLHSG SVFAGVVGVK MPRYCLFGNN   600
VTLANKFESC SVPRKINVSP TTYRLLKDCP GFVFTPRSRE ELPPNFPSEI PGICHFLDAY   660
QQGTNSKPCF QKKDVEDGNA NFLGKASGID

5

SEQ ID NO:287 PFD2 DNA SEQUENCE

Nucleic Acid Accession#:   NM_000720
Coding sequence: 119-6664 (underlined sequence corresponds to start and stop codon)

10

```
1          11         21         31         41         51
|          |          |          |          |          |
AGAATAAGGG CAGGGACCGC GGCTCCTATC TCTTGGTGAT CCCCTTCCCC ATTCCGCCCC    60
CGCCTCAACG CCCAGCACAG TGCCCTGCAC ACAGTAGTCG CTCAATAAAT GTTCGTGGAT   120
GATGATGATG ATGATGATGA AAAAAATGCA GCATCAACGG CAGCAGCAAG CGGACCACGC   180
GAACGAGGCA AACTATGCAA GAGGCACCAG ACTTCCTCTT TCTGGTGAAG GACCAACTTC   240
TCAGCCGAAT AGCTCCAAGC AAACTGTCCT GTCTTGGCAA GCTGCAATCG ATGCTGCTAG   300
ACAGGCCAAG GCTGCCCAAA CTATGAGCAC CTCTGCACCC CCACCTGTAG GATCTCTCTC   360
CCAAAGAAAA CGTCAGCAAT ACGCCAAGAG CAAAAAACAG GGTAACTCGT CCAACAGCCG   420
ACCTGCCCGC GCCCTTTTCT GTTTATCACT CAATAACCCC ATCCGAAGAG CCTGCATTAG   480
TATAGTGGAA TGGAAACCAT TTGACATATT TATATTATTG GCTATTTTTG CCAATTGTGT   540
GGCCTTAGCT ATTTACATCC CATTCCCTGA AGATGATTCT AATTCAACAA ATCATAACTT   600
GGAAAAAGTA GAATATGCCT TCCTGATTAT TTTTACAGTC GAGACATTTT TGAAGATTAT   660
AGCGTATGGA TTATTGCTAC ATCCTAATGC TTATGTTAGG AATGGATGGA ATTTACTGGA   720
TTTTGTTATA GTAATAGTAG GATTGTTTAG TGTAATTTTG GAACAATTAA CCAAAGAAAC   780
AGAAGGCGGG AACCACTCAA GCGGCAAATC TGGAGGCTTT GATGTCAAAG CCCTCCGTGC   840
CTTTCGAGTG TTGCGACCAC TTCGACTAGT GTCAGGGGTG CCCAGTTTAC AAGTTGTCCT   900
GAACTCCATT ATAAAAGCCA TGGTTCCCCT CCTTCACATA GCCCTTTTGG TATTATTTGT   960
AATCATAATC TATGCTATTA TAGGATTGGA ACTTTTTATT GGAAAAAATGC ACAAAACATG  1020
TTTTTTTGCT GACTCAGATA TCGTAGCTGA AGAGGACCCA GCTCCATGTG CGTTCTCAGG  1080
GAATGGACGC CAGTGTACTG CCAATGGCAC GGAATGTAGG AGTGGCTGGG TTGGCCCGAA  1140
CGGAGGCATC ACCAACTTTG ATAACTTTGC CTTTGCCATG CTTACTGTGT TTCAGTGCAT  1200
CACCATGGAG GGCTGGACAG ACGTGCTCTA CTGGGTAAAT GATGCGATAG GATGGGAATG  1260
GCCATGGGTG TATTTTGTTA GTCTGATCAT CCTTGGCTCA TTTTTCGTCC TTAACCTGGT  1320
TCTTGGTGTC CTTAGTGGAG AATTCTCAAA GGAAAGAGAG AAGGCAAAAG CACGGGGAGA  1380
TTTCCAGAAG CTCCGGGAGA AGCAGCAGCT GGAGGAGGAT CTAAAGGGCT ACTTGGATTG  1440
GATCACCCAA GCTGAGGACA TCGATCCGGA GAATGAGGAA GAAGGAGGAG AGGAAGGCAA  1500
ACGAAATACT AGCATGCCCA CCAGCGAGAC TGAGTCTGTG AACACAGAGA ACGTCAGCGG  1560
TGAAGGCGAG AACCGAGGCT GCTGTGGAAG TCTCTGGTGC TGGTGGAGAC GGAGAGGCGC  1620
GGCCAAGGCG GGGCCCTCTG GGTGTCGGCG GTGGGGTCAA GCCATCTCAA AATCCAAACT  1680
CAGCCGACGC TGGCGTCGCT GGAACCGATT CAATCGCAGA AGATGTAGGG CCGCCGTGAA  1740
GTCTGTCACG TTTTACTGGC TGGTTATCGT CCTGGTGTTT CTGAACACCT TAACCATTTC  1800
CTCTGAGCAC TACAATCAGC CAGATTGGTT GACACAGATT CAAGATATTG CCAACAAAGT  1860
CCTCTTGGCT CTGTTCACCT GCGAGATGCT GGTAAAAATG TACAGCTTGG GCCTCCAAGC  1920
ATATTTCGTC TCTCTTTTCA ACCGGTTTGA TTGCTTCGTG GTGTGTGGTG GAATCACTGA  1980
GACGATCCTG GTGGAACTGG AAATCATGTC TCCCCTGGGG ATCTCTGTGT TTCGGTGTGT  2040
GCGCCTCTTA AGAATCTTCA AAGTGACCAG GCACTGGACT TCCCTGAGCA ACTTAGTGGG  2100
ATCCTTATTA AACTCCATGA AGTCCATCGC TTCGCTGTTG CTTCTGCTTT TTCTCTTCAT  2160
TATCATCTTT TCCTTGCTTG GGATGCAGCT GTTTGGCGGC AAGTTTAATT TTGATGAAAC  2220
GCAAACCAAG CGGAGCACCT TTGACAATTT CCCTCAAGCA CTTCTCACAG TGTTCCAGAT  2280
CCTGACAGGC GAAGACTGGA ATGCTGTGAT GTACGATGGC ATCATGGCTT ACGGGGGCCC  2340
ATCCTCTTCA GGAATGATCG TCTGCATCTA CTTCATCATC CTCTTCATTT GTGGTAACTA  2400
TATTCTACTG AATGTCTTCT TGGCCATCGC TGTAGACAAT TTGGCTGATG CTGAAAGTCT  2460
GAACACTGCT CAGAAAGAAG AAGCGGAAGA AAAGGAGAGG AAAAAGATTG CCAGAAAAGA  2520
GAGCCTAGAA AATAAAAAGA ACAACAAACC AGAAGTCAAC CAGATAGCCA ACAGTGACAA  2580
CAAGGTTACA ATTGATGACT ATAGAGAAGA GGATGAAGAC AAGGACCCCT ATCCGCCTTG  2640
CGATGTGCCA GTAGGGGAAG AGGAAGAGGA AGAGGAGGAG GATGAACCTG AGGTTCCTGC  2700
CGGACCCCGT CCTCGAAGGA TCTCGGAGTT GAACATGAAG GAAAAAATTG CCCCCATCCC  2760
TGAAGGGAGC GCTTTCTTCA TTCTTAGCAA GACCAACCCG ATCCGCGTAG GCTGCCACAA  2820
GCTCATCAAC CACCACATCT TCACCAACCT CATCCTTGTC TTCATCATGC TGAGCAGCGC  2880
TGCCCTGGCC GCAGAGGACC CCATCCGCAG CCACTCCTTC CGGAACACGA TACTGGGTTA  2940
CTTTGACTAT GCCTTCACAG CCATCTTTAC TGTTGAGATC CTGTTGAAGA TGACAACTTT  3000
TGGAGCTTTC CTCCACAAAG GGGCCTTCTG CAGGAACTAC TTCAATTTGC TGGATATGCT  3060
GGTGGTTGGG GTGTCTCTGG TGTCATTTGG GATTCAATCC AGTGCCATCT CCGTTGTGAA  3120
GATTCTGAGG GTCTTAAGGG TCCTGCGTCC CCTCAGGGCC ATCAACAGAG CAAAAGGACT  3180
TAAGCACGTG GTCCAGTGCG TCTTCGTGGC CATCCGGACC ATCGGCAACA TCATGATCGT  3240
CACTACCCTC CTGCAGTTCA TGTTTGCCTG TATCGGGGTC CAGTTGTTCA AGGGGAAGTT  3300
CTATCGCTGT ACGGATGAAG CCAAAAGTAA CCCTGAAGAA TGCAGGGGAC TTTTCATCCT  3360
CTACAAGGAT GGGGATGTTG ACAGTCCTGT GGTCCGTGAA CGGATCTGGC AAAACAGTGA  3420
TTTCAACTTC GACAACGTCC TCTCTGCTAT GATGGCGCTC TTCACAGTCT CCACGTTTGA  3480
GGGCTGGCCT GCGTTGCTGT ATAAAGCCAT CGACTCGAAT GGAGAGAACA TCGGCCCAAT  3540
CTACAACCAC CGCGTGGAGA TCTCCATCTT CTTCATCATC TACATCATCA TTGTAGCTTT  3600
CTTCATGATG AACATCTTTG TGGGCTTTGT CATCGTTACA TTTCAGGAAC AAGGAGAAAA  3660
AGAGTATAAG AACTGTGAGC TGGACAAAAA TCAGCGTCAG TGTGTTGAAT ACGCCTTGAA  3720
AGCACGTCCC TTGCGGAGAT ACATCCCCAA AAACCCCTAC CAGTACAAGT TCTGGTACGT  3780
GGTGAACTCT TCGCCTTTCG AATACATGAT GTTTGTCCTC ATCATGCTCA ACACACTCTG  3840
CTTGGCCATG CAGCACTACG AGCAGTCCAA GATGTTCAAT GATGCCATGG ACATTCTGAA  3900
CATGGTCTTC ACCGGGGTGT CACCGTCGA GATGGTTTTG AAAGTCATCG CATTTAAGCC  3960
TAAGGGGTAT TTTAGTGACG CCTGGAACAC GTTTGACTCC CTCATCGTAA TCGGCAGCAT  4020
TATAGACGTG GCCCTCAGCG AAGCGGACCC AACTGAAAGT GAAAATGTCC CTGTCCCAAC  4080
```

422

```
     TGCTACACCT GGGAACTCTG AAGAGAGCAA TAGAATCTCC ATCACCTTTT TCCGTCTTTT  4140
     CCGAGTGATG CGATTGGTGA AGCTTCTCAG CAGGGGGGAA GGCATCCGGA CATTGCTGTG  4200
     GACTTTTATT AAGTCCTTTC AGGCGCTCCC GTATGTGGCC CTCCTCATAG CCATGCTGTT  4260
     CTTCATCTAT GCCGGTCATTG GCATGCAGAT GTTTGGGAAA GTTGCCATGA GAGATAACAA  4320
     CCAGATCAAT AGGAACAATA ACTTCCAGAC GTTTCCCCAG GCGGTGCTGC TGCTCTTCAG  4380
     GTGTGCAACA GGTGAGGCCT GGCAGGAGAT CATGCTGGCC TGTCTCCCAG GGAAGCTCTG  4440
     TGACCCTGAG TCAGATTACA ACCCCGGGGA GGAGTATACA TGTGGGAGCA ACTTTGCCAT  4500
     TGTCTATTTC ATCAGTTTTT ACATGCTCTG TGCATTTCTG ATCATCAATC TGTTTGTGGC  4560
     TGTCATCATG GATAATTTCG ACTATCTGAC CCGGGACTGG TCTATTTTGG GGCCTCACCA  4620
     TTTAGATGAA TTCAAAAGAA TATGGTCAGA ATATGACCCT GAGGCAAAGG GAAGGATAAA  4680
     ACACCTTGAT GTGGTCACTC TGCTTCGACG CATCCAGCCT CCCCTGGGGT TTGGGAAGTT  4740
     ATGTCCACAC AGGGTAGCGT GCAAGAGATT AGTTGCCATG AACATGCCTC TCAACAGTGA  4800
     CGGGACAGTC ATGTTTAATG CAACCCTGTT TGCTTTGGTT CGAACGGCTC TTAAGATCAA  4860
     GACCGAAGGG AACCTGGAGC AAGCTAATGA AGAACTTCGG GCTGTGATAA AGAAAATTTG  4920
     GAAGAAAACC AGCATGAAAT TACTTGACCA AGTTGTCCCT CCAGCTGGTG ATGATGAGGT  4980
     AACCGTGGGG AAGTTCTATG CCACTTTCCT GATACAGGAC TACTTTAGGA AATTCAAGAA  5040
     ACGGAAAGAA CAAGGACTGG TGGGAAAGTA CCCTGCGAAG AACACCACAA TTGCCCTACA  5100
     GGCGGGATTA AGGACACTGC ATGACATTGG GCCAGAAATC CGGCGTGCTA TATCGTGTGA  5160
     TTTGCAAGAT GACGAGCCTG AGGAAACAAA ACGAGAAGAA GAAGATGATG TGTTCAAAAG  5220
     AAATGGTGCC CTGCTTGGAA ACCATGTCAA TCATGTTAAT AGTGATAGGA GAGATTCCCT  5280
     TCAGCAGACC AATACCACCC ACCGTCCCCT GCATGTCCAA AGGCCTTCAA TTCCACCTGC  5340
     AAGTGATACT GAGAAACCGC TGTTTCCTCC AGCAGGAAAT TCGGTGTGTC ATAACCATCA  5400
     TAACCATAAT TCCATAGGAA AGCAAGTTCC CACCTCAACA ATGCCAATC TCAATAATGC  5460
     CAATATGTCC AAAGCTGCCC ATGGAAAGCG GCCCAGCATT GGGAACCTTG AGCATGTGTC  5520
     TGAAAATGGG CATCATTCTT CCCACAAGCA TGACCGGGAG CCTCAGAGAA GGTCCAGTGT  5580
     GAAAAGAACC CGCTATTATG AAACTTACAT TAGGTCCGAC TCAGGAGATG AACAGCTCCC  5640
     AACTATTTGC CGGGAAGACC CAGAGATACA TGGCTATTTC AGGGACCCCC ACTGCTTGGG  5700
     GGAGCAGGAG TATTTCAGTA GTGAGGAATG CTACGAGGAT GACAGCTCGC CCACCTGGAG  5760
     CAGGCAAAAC TATGGCTACT ACAGCAGATA CCCCAGGCAGA AACATCGACT CTGAGAGGCC  5820
     CCGAGGCTAC CATCATCCCC AAGGATTCTT GGAGGACGAT GACTCGCCCG TTTGCTATGA  5880
     TTCACGGAGA TCTCCAAGGA GACGCCTACT ACCTCCCACC CCAGCATCCC ACCGGAGATC  5940
     CTCCTTCAAC TTTGAGTGCC TGCGCCGGCA GAGCAGCCAG GAAGAGGTCC CGTCGTCTCC  6000
     CATCTTCCCC CATCGCACGG CCCTGCCTCT GCATCTAATG CAGCAACAGA TCATGGCAGT  6060
     TGCCGGCCTA GATTCAAGTA AAGCCCAGAA GTACTCACCG AGTCACTCGA CCCGGTCGTG  6120
     GGCCACCCCT CCAGCAACCC CTCCCTACCG GGACTGGACA CCGTGCTACA CCCCCCTGAT  6180
     CCAAGTGGAG CAGTCAGAGG CCCTGGACCA GGTGAACGGC AGCCTGCCGT CCCTGCACCG  6240
     CAGCTCCTGG TACACAGACG AGCCCGACAT CTCCTACCGG ACTTTCACAC CAGCCAGCCT  6300
     GACTGTCCCC AGCAGCTTCC GGAACAAAAA CAGCGACAAG CAGAGGAGTG CGGACAGCTT  6360
     GGTGGAGGCA GTCCTGATAT CCGAAGGCTT GGGACGCTAT GCAAGGGACC CAAAATTTGT  6420
     GTCAGCAACA AAACACGAAA TCGCTGATGC CTGTGACCTC ACCATCGACG AGATGGAGAG  6480
     TGCAGCCAGC ACCCTGCTTA ATGGGAACGT GCGTCCCCGA GCCAACGGGG ATGTGGGCCC  6540
     CCTCTCACAC CGGCAGGACT ATGAGCTACA GGACTTTGGT CCTGGCTACA GCGACGAAGA  6600
     GCCAGACCCT GGGAGGGATG AGGAGGACCT GGCGGATGAA ATGATATGCA TCACCACCTT  6660
     GTAGCCCCCA GCGAGGGGCA GACTGGCTCT GGCCTCAGGT GGGGCGCAGG AGAGCCAGGG  6720
     GAAAAGTGCC TCATAGTTAG GAAAGTTTAG GCACTAGTTG GGAGTAATAT TCAATTAATT  6780
     AGACTTTTGT ATAAGAGATG TCATGCCTCA AGAAAGCCAT AAACCTGGTA GGAACAGGTC  6840
     CCAAGCGGTT GAGCCTGGCA GAGTACCATG CGCTCGGCCC CAGCTGCAGG AAACAGCAGG  6900
     CCCCGCCCTC TCACAGAGGA TGGGTGAGGA GGCCAGACCT GCCCTGCCCC ATTGTCCAGA  6960
     TGGGCACTGC TGTGGAGTCT GCTTCTCCCA TGTACCAGGG CACCAGGCCC ACCCAACTGA  7020
     AGGCATGGCG GCGGGGTGCA GGGGAAAGTT AAAGGTGATG ACGATCATCA CACCTCGTGT  7080
     CGTTACCTCA GCCATCGGTC TAGCATATCA GTCACTGGGC CCAACATATC CATTTTTAAA  7140
     CCCTTTCCCC CAAATACACT GCGTCCTGGT TCCTGTTTAG CTGTTCTGAA ATA
```

SEQ ID NO:288 PFD2 Protein sequence:
Protein Accession #:   A38198

```
     1          11         21         31         41         51
     |          |          |          |          |          |
     MMMMMMMKKM QHQRQQQADH ANEANYARGT RLPLSGEGPT SQPNSSKQTV LSWQAAIDAA   60
     RQAKAAQTMS TSAPPPVGSL SQRKRQQYAK SKKQGNSSNS RPARALFCLS LNNPIRRACI  120
     SIVEWKPFDI FILLAIFANC VALAIYIPPP EDDSNSTNHN LEKVEYAFLI IFTVETFLKI  180
     IAYGLLLHPN AYVRNGWNLL DFVIVIVGLF SVILEQLTKE TEGGNHSSGK SGGFDVKALR  240
     AFRVLRPLRL VSGVPSLQVV LNSIIKAMVP LLHIALLVLF VIIIYAIIGL ELFIGKMHKT  300
     CFFADSDIVA EEDPAPCAFS GNGRQCTANG TECRSGWVGP NGGITNFDNF AFAMLTVFQC  360
     ITMEGWTDVL YWVNDAIGWE WPWVYFVSLI ILGSFFVLNL VLGVLSGEFS KEREKAKARG  420
     DFQKLREKQQ LEEDLKGYLD WITQAEDIDP ENEEEGGEEG KRNTSMPTSE TESVNTENVS  480
     GEGENRGCCG SLWCWWRRRG AAKAGPSGCR RWGQAISKSK LSRRWRRWNR FNRRRCRAAV  540
     KSVTFYWLVI VLVFLNTLTI SSEHYNQPDW LTQIQDIANK VLLALFTCEM LVKMYSLGLQ  600
     AYFVSLFNRF DCFVVCGGIT ETILVELEIM SPLGISVFRC VRLLRIFKVT RHWTSLSNLV  660
     ASLLNSMKSI ASLLLLLPLF IIIFSLLGMQ LFGGKFNFDE TQTKRSTFDN FPQALLTVFQ  720
     ILTGEDWNAV MYDGIMAYGG PSSSGMIVCI YFIILFICGN YILLNVFLAI AVDNLADAES  780
     LNTAQKEEAE EKERKKIARK ESLENKKNNK PEVNQIANSD NKVTIDDYRE EDEDKDPYPP  840
     CDVPVGEEEE EEEEDEPEVP AGPRPRRISE LNMKEKIAPI PEGSAFFILS KTNPIRVGCH  900
     KLINHHIFTN LILVFIMLSS AALAAEDPIR SHSFRNTILG YFDYAFTAIF TVEILLKMTT  960
     FGAFLHKGAF CRNYFNLLDM LVVGVSLVSF GIQSSAISVV KILRVLRVLR PLRAINRAKG 1020
     LKHVVQCVFV AIRTIGNIMI VTTLLQFMFA CIGVQLPKGK FYRCTDEAKS NPEECRGLFI 1080
     LYKDGDVDSP VVRERIWQNS DFNFDNVLSA MMALFTVSTF EGWPALLYKA IDSNGENIGP 1140
     IYNHRVEISI FFIIYIIIVA FFMMNIFVGF VIVTFQEQGE KEYKNCELDK NQRQCVEYAL 1200
     KARPLRRYIP KNPYQYKFWY VVNSSPFEYM MFVLIMLNTL CLAMQHYEQS KMFNDAMDIL 1260
     NMVFTGVFTV EMVLKVIAFK PKGYFSDAWN TFDSLIVIGS IIDVALSEAD PTESENVPVP 1320
```

423

```
     TATPGNSEES NRISITFFRL FRVMRLVKLL SRGEGIRTLL WTFIKSFQAL PYVALLIAML   1380
     FFIYAVIGMQ MFGKVAMRDN NQINRNNNFQ TFPQAVLLLF RCATGEAWQE IMLACLPGKL   1440
     CDPESDYNPG EEYTCGSNFA IVYFISFYML CAFLIINLFV AVIMDNFDYL TRDWSILGPH   1500
     HLDEFKRIWS EYDPEAKGRI KHLDVVTLLR RIQPPLGFGK LCPHRVACKR LVAMNMPLNS   1560
     DGTVMFNATL FALVRTALKI KTEGNLEQAN EELRAVIKKI WKKTSMKLLD QVVPPAGDDE   1620
     VTVGKFYATF LIQDYFRKFK KRKEQGLVGK YPAKNTTIAL QAGLRTLHDI GPEIRRAISC   1680
     DLQDDEPEET KREEEDDVFK RNGALLGNHV NHVNSDRRDS LQQTNTTHRP LHVQRPSIPP   1740
     ASDTEKPLFP PAGNSVCHNH HNHNSIGKQV PTSTNANLNN ANMSKAAHGK RPSIGNLEHV   1800
     SENGHHSSHK HDREPQRRSS VKRTRYYETY IRSDSGDEQL PTICREDPEI HGYFRDPHCL   1860
     GEQEYFSSEE CYEDDSSPTW SRQNYGYYSR YPGRNIDSER PRGYHHPQGF LEDDDSPVCY   1920
     DSRRSPRRRL LPPTPASHRR SSFNFECLRR QSSQEEVPSS PIFPHRTALP LHLMQQQIMA   1980
     VAGLDSSKAQ KYSPSHSTRS WATPPATPPY RDWTPCYTPL IQVEQSEALD QVNGSLPSLH   2040
     RSSWYTDEPD ISYRTFTPAS LTVPSSFRNK NSDKQRSADS LVEAVLISEG LGRYARDPKF   2100
     VSATKHEIAD ACDLTIDEME SAASTLLNGN VRPRANGDVG PLSHRQDYEL QDFGPGYSDE   2160
     EPDPGRDEED LADEMICITT L
```

SEQ ID NO:289 OBI6 DNA SEQUENCE

Nucleic Acid Accession#: NM_002812

Coding sequence: 150-3362 (underlined sequence corresponds to start and stop codon)

```
     1          11         21         31         41         51
     |          |          |          |          |          |
     AACTCCCGCC TCGGGACGCC TCGGGGTCGG GCTCCGGCTG CGGCTGCTGC TGCGGCGCCC    60
     GCGCTCCGGT GCGTCCGCCT CCTGTGCCCG CCGCGGAGCA GTCTGCGGCC CGCCGTGCGC   120
     CCTCAGCTCC TTTTCCTGAG CCCGCCGCGA TGGGAGCTGC GCGGGGATCC CCGGCCAGAC   180
     CCCGCCGGTT GCCTCTGCTC AGCGTCCTGC TGCTGCCGCT GCTGGGCGGT ACCCAGACAG   240
     CCATTGTCTT CATCAAGCAG CCGTCCTCCC AGGATGCACT GCAGGGGCGC CGGGCGCTGC   300
     TTCGCTGTGA GGTTGAGGCT CCGGGGCCGG TACATGTGTA CTGGCTGCTC GATGGGGCCC   360
     CTGTCCAGGA CACGGAGCGG CGTTTCGCCC AGGGCAGCAG CCTGAGCTTT GCAGCTGTGG   420
     ACCGGCTGCA GGACTCTGGC ACCTTCCAGT GTGTGGCTCG GGATGATGTC ACTGGAGAAG   480
     AAGCCCGCAG TGCCAACGCC TCCTTCAACA TCAAATGGAT TGAGGCAGGT CCTGTGGTCC   540
     TGAAGCATCC AGCCTCGGAA GCTGAGATCC AGCCACAGAC CCAGGTCACA CTTCGTTGCC   600
     ACATTGATGG GCACCCTCGG CCCACCTACC AATGGTTCCG AGATGGGACC CCCCTTTCTG   660
     ATGGTCAGAG CAACCACACA GTCAGCAGCA AGGAGCGGAA CCTGACGCTC CGGCCAGCTG   720
     GTCCTGAGCA TAGTGGGCTG TATTCCTGCT GCGCCCACAG TGCTTTTGGC CAGGCTTGCA   780
     GCAGCCAGAA CTTCACCTTG AGCATTGCTG ATGAAAGCTT TGCCAGGGTG GTGCTGGCAC   840
     CCCAGGACGT GGTAGTAGCG AGGTATGAGG AGGCCATGTT CCATTGCCAG TTCTCAGCCC   900
     AGCCACCCCC GAGCCTGCAG TGGCTCTTTG AGGATGAGAC TCCCATCACT AACCGCAGTG   960
     GCCCCCCACA CCTCCGCAGA GCCACAGTGT TTGCCAACGG GTCTCTGCTG CTGACCCAGG   1020
     TCCGGCCACG CAATGCAGGG ATCTACCGCT GCATTGGCCA GGGGCAGAGG GGCCCACCCA   1080
     TCATCCTGGA AGCCACACTT CACCTAGCAG AGATTGAAGA CATGCCGCTA TTTGAGCCAC   1140
     GGGTGTTTAC AGCTGGCAGC GAGGAGCGTG TGACCTGCCT TCCCCCCAAG GGTCTGCCAG   1200
     AGCCCAGCGT GTGGTGGGAG CACGCGGGAG TCCGGCTGCC CACCCATGGC AGGGTCTACC   1260
     AGAAGGGCCA CGAGCTGGTG TTGGCCAATA TTGCTGAAAG TGATGCTGGT GTCTACACCT   1320
     GCCACGCGGC CAACCTGGCT GGTCAGCGGA GACAGGATGT CAACATCACT GTGGCCACTG   1380
     TGCCCTCCTG GCTGAAGAAG CCCCAAGACA GCCAGCTGGA GGAGGGCAAA CCCGGCTACT   1440
     TGGATTGCCT GACCCAGGCC ACACCAAAAC CTACAGTTGT CTGGTACAGA AACCAGATGC   1500
     TCATCTCAGA GGACTCACGG TTCGAGGTCT TCAAGAATGG GACCTTGCGC ATCAACAGCG   1560
     TGGAGGTGTA TGATGGGACA TGGTACCGTT GTATGAGCAG CACCCCAGCC GGCAGCATCG   1620
     AGGCGCAAGC CCGTGTCCAA GTGCTGGAAA AGCTCAAGTT CACACCACCA CCCCAGCCAC   1680
     AGCAGTGCAT GGAGTTTGAC AAGGAGGCCA CGGTGCCCTG TTCAGCCACA GGCGAGAGA   1740
     AGCCCACTAT TAAGTGGGAA CGGGCAGATG GGAGCAGCCT CCCAGAGTGG GTGACAGACA   1800
     ACGCTGGGAC CCTGCATTTT GCCCGGGTGA CTCGAGATGA CGCTGGCAAC TACACTTGCA   1860
     TTGCCTCCAA CGGGCCGCAG GGCCAGATTC GTGCCCATGT CCAGCTCACT GTGGCAGTTT   1920
     TTATCACCTT CAAAGTGGAA CCAGAGCGTA CGACTGTGTA CCAGGGCCAC ACAGCCCTAC   1980
     TGCAGTGCGA GGCCCAGGGG GACCCCAAGC CGCTGATTCA GTGGAAAGGC AAGGACCGCA   2040
     TCCTGGACCC CACCAAGCTG GGACCCAGGA TGCACATCTT CCAGAATGGC TCCCTGGTGA   2100
     TCCATGACGT GGCCCCTGAG GACTCAGGCC GCTACACCTG CATTGCAGGC AACAGCTGCA   2160
     ACATCAAGCA CACGGAGGCC CCCCTCTATG TCGTGGACAA GCCTGTGCCG GAGGAGTCGG   2220
     AGGGCCCTGG CAGCCCTCCC CCCTACAAGA TGATCCAGAC CATTGGGTTG TCGGTGGGTG   2280
     CCGCTGGTGC CTACATCATT GCCGTGCTGG GCCTCATGTT CTACTGCAAG AAGCGCTGCA   2340
     AAGCCAAGCG GCTGCAGAAG CAGCCCGAGG GCGAGGAGCC AGAGATGGAA TGCCTCAACG   2400
     GAGGGCCTTT GCAGAACGGG CAGCCCTCAG CAGAGATCCA AGAAGAAGTG GCCTTGACCA   2460
     GCTTGGGCTC CGGCCCCGCG GCCACCAACA AACGCCACAG CACAAGTGAT AAGATGCACT   2520
     TCCCACGGTC TAGCCTGCAG CCCATCACCA CGCTGGGGAA GAGTGAGTTT GGGGAGGTGT   2580
     TCCTGGCAAA GGCTCAGGGC TTGGAGGAGG GAGTGGCAGA GACCCTGGTA CTTGTGAAGA   2640
     GCCTGCAGAC GAAGGATGAG CAGCAGCAGC TGGACTTCCG GAGGGGAGTTG GAGATGTTTG   2700
     GGAAGCTGAA CCACGCCAAC GTGGTGCGGC TCCTGGGGGT GTGCCGGGAG GCTGAGCCCC   2760
     ACTACATGGT GCTGGAATAT GTGGATCTGG GAGACCTCAA GCAGTTCCTG AGGATTCCA   2820
     AGAGCAAGGA TGAAAAATTG AAGTCACAGC CCCTCAGCAC CAAGCAGAAG GTGGCCCTAT   2880
     GCACCCAGGT AGCCCTGGGC ATGGAGCACC TGTCCAACAA CCGCTTTGTG CATAAGGACT   2940
     TGGCTGCGCG TAACTGCCTG GTCAGTGCCC AGAGACAAGT GAAGGTGTCT GCCCTGGGCC   3000
     TCAGCAAGGA TGTGTACAAC AGTGAGTACT ACCACTTCCG CCAGGCCTGG GTGCCGCTGC   3060
     GCTGGATGTC CCCCGAGGCC ATCCTGGAGG GTGACTTCTC TACCAAGTCT GATGTCTGGG   3120
     CCTTCGGTGT GCTGATGTGG GAAGTGTTTA CACATGGAGA GATGCCCCAT GGTGGGCAGG   3180
     CAGATGATGA AGTACTGGCA GATTTGCAGG CTGGGAAGGC TAGACTTCCT CAGCCCGAGG   3240
     GCTGCCCTTC CAAACTCTAT CGGCTGATGC AGCGCTGCTG GGCCCTCAGC CCCAAGGACC   3300
     GGCCCTCCTT CAGTGAGATT GCCAGCGCCC TGGGAGACAG CACCGTGGAC AGCAAGCCGT   3360
     GAGGAGGGGAG CCCGCTCAGG ATGGCCTGGG CAGGGGAGGA CATCTCTAGA GGGAAGCTCA   3420
```

424

```
        CAGCATGATG GGCAAGATCC CTGTCCTCCT GGGCCCTGAG GTGCCCTAGT GCAACAGGCA   3480
        TTGCTGAGGT CTGAGCAGGG CCTGGCCTTT CCTCCTCTTC CTCACCCTCA TCCTTTGGGA   3540
        GGCTGACTTG GACCCAAACT GGGCGACTAG GGCTTTGAGC TGGGCAGTTT CCCCTGCCAC   3600
        CTCTTCCTCT ATCAGGGACA GTGTGGGTGC CACAGGTAAC CCCAATTTCT GGCCTTCAAC   3660
        TTCTCCCCTT GACCGGGTCC AACTCTGCCA CTCATCTGCC AACTTTGCCT GGGGAGGGCT   3720
        AGGCTTGGGA TGAGCTGGGT TTGTGGGGAG TTCCTTAATA TTCTCAAGTT CTGGGCACAC   3780
        AGGGTTAATG AGTCTCTTGC CCACTGGTCC ACTTGGGGGT CTAGACCAGG ATTATAGAGG   3840
        ACACAGCAAG TGAGTCCTCC CCACTCTGGG CTTGTGCACA CTGACCCAGA CCCACGTCTT   3900
        CCCCACCCTT CTCTCCTTTC CTCATCCTAA GTGCCTGGCA GATGAAGGAG TTTTCAGGAG   3960
        CTTTTGACAC TATATAAACC GCCCTTTTTG TATGCACCAC GGGCGGCTTT TATATGTAAT   4020
        TGCAGCGTGG GGTGGGTGGG CATGGGAGGT AGGGGTGGGC CCTGGAGATG AGGAGGGTGG   4080
        GCCATCCTTA CCCCACACTT TTATTGTTGT CGTTTTTTGT TTGTTTTGTT TTTTTGTTTT   4140
        TGTTTTTGTT TTTACACTCG CTGCTCTCAA TAAATAAGCC TTTTTTA
```

SEQ ID NO:290 <u>QBI6 Protein sequence:</u>
Protein Accession #: NP_002812

```
        1          11         21         31         41         51
        |          |          |          |          |          |
        MGAARGSPAR PRRLPLLSVL LLPLLGGTQT AIVFIKQPSS QDALQGRRAL LRCEVEAPGP     60
        VHVYWLLDGA PVQDTERRFA QGSSLSFAAV DRLQDSGTFQ CVARDDVTGE EARSANASFN    120
        IKWIEAGPVV LKHPASEAEI QPQTQVTLRC HIDGHPRPTY QWFRDGTPLS DGQSNHTVSS    180
        KERNLTLRPA GPEHSGLYSC CAHSAFGQAC SSQNFTLSIA DESFARVVLA PQDVVVARYE    240
        EAMFHCQFSA QPPPSLQWLF EDETPITNRS RPPHLRRATV FANGSLLLTQ VRPRNAGIYR    300
        CIGQGQRGPP IILEATLHLA EIEDMPLFEP RVFTAGSEER VTCLPPKGLP EPSVWWEHAG    360
        VRLPTHGRVY QKGHELVLAN IAESDAGVYT CHAANLAGQR RQDVNITVAT VPSWLKKPQD    420
        SQLEEGKPGY LDCLTQATPK PTVVWYRNQM LISEDSRFEV FKNGTLRINS VEVYDGTWYR    480
        CMSSTPAGSI EAQARVQVLE KLKFTPPPQP QQCMEFDKEA TVPCSATGRE KPTIKWERAD    540
        GSSLPEWVTD NAGTLHFARV TRDDAGNYTC IASNGPQGQI RAHVQLTVAV PITFKVEPER    600
        TTVYQGHTAL LQCEAQGDPK PLIQWKGKDR ILDPTKLGPR MHIFQNGSLV IHDVAPEDSG    660
        RYTCIAGNSC NIKHTEAPLY VVDKPVPEES EGPGSPPPYK MIQTIGLSVG AAVAYIIAVL    720
        GLMFYCKKRC KAKRLQKQPE GEEPEMECLN GGPLQNGQPS AEIQEEVALT SLGSGPAATN    780
        KRHSTSDKMH FPRSSLQPIT TLGKSEFGEV FLAKAQGLEE GVAETLVLVK SLQTKDEQQQ    840
        LDFRRELEMF GKLNHANVVR LLGLCREAEP HYMVLEYVDL GDLKQFLRIS KSKDEKLKSQ    900
        PLSTKQKVAL CTQVALGMEH LSNNRFVHKD LAARNCLVSA QRQVKVSALG LSKDVYNSEY    960
        YHFRQAWVPL RWMSPEAILE GDFSTKSDVW AFGVLMWEVF THGEMPHGGQ ADDEVLADLQ   1020
        AGKARLPQPE GCPSKLYRLM QRCWALSPKD RPSFSEIASA LGDSTVDSKP
```

SEQ ID NO:291 AAB1 DNA SEQUENCE

Nucleic Acid Accession #: NM_002205
Coding sequence: 1-3150 (underlined sequences correspond to start and stop codons)

```
        1          11         21         31         41         51
        |          |          |          |          |          |
        ATGGGGAGCC GGACGCCAGA GTCCCCTCTC CACGCCGTGC AGCTGCGCTG GGGCCCCCGG     60
        CGCCGACCCC CGCTSSTGCC GCTGCTGTTG CTGCTSSTGC CGCCGCCACC CAGGGTCGGG    120
        GGCTTCAACT TAGACGCGGA GGCCCCAGCA GTACTCTCGG GGCCCCCGGG CTCCTTCTTC    180
        GGATTCTCAG TGGAGTTTTA CCGGCCGGGA ACAGACGGGG TCAGTGTGCT GGTGGGAGCA    240
        CCCAAGGCTA ATACCAGCCA GCCAGGAGTG CTGCAGGGTG GTGCTGTCTA CCTCTGTCCT    300
        TGGGGTGCCA GCCCCACACA GTGCACCCCC ATTGAATTTG ACAGCAAAGG CTCTCGGCTC    360
        CTGGAGTCCT CACTGTCCAG CTCAGAGGGA GAGGAGCCTG TGGAGTACAA GTCCTTGCAG    420
        TGGTTCGGGG CAACAGTTCG AGCCCATGGC TCCTCCATCT TGGCATGCGC TCCACTGTAC    480
        AGCTGGCGCA CAGAGAAGGA GCCACTGAGC GACCCCGTGG GCACCTGCTA CCTCTCCACA    540
        GATAACTTCA CCCGAATTCT GGAGTATGCA CCCTGCCGCT CAGATTTCAG CTGGGCAGCA    600
        GGACAGGGTT ACTGCCAAGG AGGCTTCAGT GCCGAGTTCA CCAAGACTGG CCGTGTGGTT    660
        TTAGGTGGAC CAGGAAGCTA TTTCTGGCAA GGCCAGATCC TGTCTGCCAC TCAGGAGCAG    720
        ATTGCAGAAT CTTATTACCC CGAGTACCTG ATCAACCTGG TTCAGGGGCA GCTGCAGACT    780
        CGCCAGGCCA GTTCCATCTA TGATGACAGC TACCTAGGAT ACTCTGTGGC TGTTGGTGAA    840
        TTCAGTGGTG ATGACACAGA AGACTTTGTT GCTGGTGTGC CCAAAGGGAA CCTCACTTAC    900
        GGCTATGTCA CCATCCTTAA TGGCTCAGAC ATTCGATCCC TCTACAACTT CTCAGGGGAA    960
        CAGATGGCCT CCTACTTTGG CTATGCAGTG GCCGCCACAG ACGTCAATGG GGACGGGCTG   1020
        GATGACTTGC TGGTGGGGGC ACCCCTGCTC ATGGATCGGA CCCCTGACGG GCGGCCTCAG   1080
        GAGGTGGGCA GGGTCTACGT CTACCTGCAG CACCCAGCCG GCATAGAGCC CACGCCCACC   1140
        CTTACCCTCA CTGGCCATGA TGAGTTTGGC CGATTTGGCA GCTCCTTGAC CCCCCTGGGG   1200
        GACCTGGACC AGGATGGCTA CAATGATGTG GCCATCGGGG CTCCCTTTGG TGGGGAGACC   1260
        CAGCAGGGAG TAGTGTTTGT ATTTCCTGGG GGCCCAGGAG GGCTGGGCTC TAAGCCTTCC   1320
        CAGGTTCTGC AGCCCCTGTG GGCAGCCAGC CACACCCCAG ACTTCTTTGG CTCTGCCCTT   1380
        CGAGGAGGCC GAGACCTGGA TGGCAATGGA TATCCTGATC TGATTGTGGG GTCCTTTGGT   1440
        GTGGACAAGG CTGTGGTATA CAGGGGCCGC CCCATCGTGT CCGCTAGTGC CTCCCTCACC   1500
        ATCTTCCCCG CCATGTTCAA CCCAGAGGAG CGGAGCTGCA GCTTAGAGGG GAACCCTGTG   1560
        GCCTGCATCA ACCTTAGCTT CTGCCTCAAT GCTTCTGGAA AACACGTTGC TGACTCCATT   1620
        GGTTTCACAG TGGAACTTCA GCTGGACTGG CAGAAGCAGA AGGGAGGGGT ACGGCGGGCA   1680
        CTGTTCCTGG CCTCCAGGCA GGCAACCCTG ACCCAGACCC TGCTCATCCA GAATGGGGCT   1740
        CGAGAGGATT GCAGAGAGAT GAAGATCTAC CTCAGGAACG AGTCAGAATT TCGAGACAAA   1800
        CTCTCGCCGA TTCACATCGC TCTCAACTTC TCCTTGGACC CCCAAGCCCC AGTGGACAGC   1860
        CACGGCCTCA GGCCAGCCCT ACATTATCAG AGCAAGAGCC GGATAGAGGA CAAGGCTCAG   1920
        ATCTTGCTGG ACTGTGGAGA AGACAACATC TGTGTGCCTG ACCTGCAGCT GGAAGTGTTT   1980
```

425

```
GGGGAGCAGA ACCATGTGTA CCTGGGTGAC AAGAATGCCC TGAACCTCAC TTTCCATGCC  2040
CAGAATGTGG GTGAGGGTGG CGCCTATGAG GCTGAGCTTC GGGTCACCGC CCCTCCAGAG  2100
GCTGAGTACT CAGGACTCGT CAGACACCCA GGGAACTTCT CCAGCCTGAG CTGTGACTAC  2160
TTTGCCGTGA ACCAGAGCCG CCTGCTGGTG TGTGACCTGG GCAACCCCAT GAAGGCAGGA  2220
GCCAGTCTGT GGGGTGGCCT TCGGTTTACA GTCCCTCATC TCCGGGACAC TAAGAAAACC  2280
ATCCAGTTTG ACTTCCAGAT CCTCAGCAAG AATCTCAACA ACTCGCAAAG CGACGTGGTT  2340
TCCTTTCGGC TCTCCGTGGA GGCTCAGGCC CAGGTCACCC TGAACGGTGT CTCCAAGCCT  2400
GAGGCAGTGC TATTCCCAGT AAGCGACTGG CATCCCCGAG ACCAGCCTCA GAAGGAGGAG  2460
GACCTGGGAC CTGCTGTCCA CCATGTCTAT GAGCTCATCA ACCAAGGCCC CAGCTCCATT  2520
AGCCAGGGTG TGCTGGAACT CAGCTGTCCC CAGGCTCTGG AAGGTCAGCA GCTCCTATAT  2580
GTGACCAGAG TTACGGGACT CAACTGCACC ACCAATCACC CCATTAACCC AAAGGGCCTG  2640
GAGTTGGATC CCGAGGGTTC CCTGCACCAC CAGCAAAAAC GGGAAGCTCC AAGCCGCAGC  2700
TCTGCTTCCT CGGGACCTCA GATCCTGAAA TGCCCGGAGG CTGAGTGTTT CAGGCTGCGC  2760
TGTGAGCTCG GGCCCCTGCA CCAACAAGAG AGCCAAAGTC TGCAGTTGCA TTTCCGAGTC  2820
TGGGCCAAGA CTTTCTTGCA GCGGGAGCAC CAGCCATTTA GCCTGCAGTG TGAGGCTGTG  2880
TACAAAGCCC TGAAGATGCC CTACCGAATC CTGCCTCGGC AGCTGCCCCA AAAAGAGCGT  2940
CAGGTGGCCA CAGCTGTGCA ATGGACCAAG GCAGAAGGCA GCTATGGCGT CCCACTGTGG  3000
ATCATCATCC TAGCCATCCT GTTTGGCCTC CTGCTCCTAG GTCTACTCAT CTACATCCTC  3060
TACAAGCTTG GATTCTTCAA ACGCTCCCTC CCATATGGCA CCGCCATGGA AAAAGCTCAG  3120
CTCAAGCCTC CAGCCACCTC TGATGCCTGA
```

SEQ ID NO:292 AAB1 Protein sequence:
Protein Accession #:   NP_002196

```
1          11         21         31         41         51
|          |          |          |          |          |
MGSRTPESPL HAVQLRWGPR RRPPLLPLLL LLLPPPPRVG GFNLDAEAPA VLSGPPGSFF   60
GFSVEFYRPG TDGVSVLVGA PKANTSQPGV LQGGAVYLCP WGASPTQCTP IEFDSKGSRL  120
LESSLSSSEG EEPVEYKSLQ WFGATVRAHG SSILACAPLY SWRTEKEPLS DPVGTCYLST  180
DNFTRILEYA PCRSDFSWAA GQGYCQGGFS AEFTKTGRVV LGGPGSYFWQ GQILSATQEQ  240
IAESYYPEYL INLVQGQLQT RQASSIYDDS YLGYSVAVGE FSGDDTEDFV AGVPKGNLTY  300
GYVTILNGSD IRSLYNFSGE QMASYFGYAV AATDVNGDGL DDLLVGAPLL MDRTPDGRPQ  360
EVGRVVVYLQ HPAGIEPTPT LTLTGHDEFG RFGSSLTPLG DLDQDGYNDV AIGAPFGGET  420
QQGVVFVFPG GPGGLGSKPS QVLQPLWAAS HTPDFFGSAL RGGRDLDGNG YPDLIVGSFG  480
VDKAVVYRGR PIVSASASLT IFFPAMFNPEE RSCSLEGNPV ACINLSFCLN ASGKHVADSI  540
GFTVELQLDW QKQKGGVRRA LFLASRQATL TQTLLIQNGA REDCREMKIY LRNESEFRDK  600
LSPIHIALNF SLDPQAPVDS HGLRPALHYQ SKSRIEDKAQ ILLDCGEDNI CVPDLQLEVF  660
GEQNHVYLGD KNALNLTFHA QNVGEGGAYE AELRVTAPPE AEYSGLVRHP GNFSSLSCDY  720
FAVNQSRLLV CDLGNPMKAG ASLWGGLRFT VPHLRDTKKT IQFDFQILSK NLNNSQSDVV  780
SFRLSVEAQA QVTLNGVSKP EAVLFPVSDW HPRDQPQKEE DLGPAVHHVY ELINQGPSSI  840
SQGVLELSCP QALEGQQLLY VTRVTGLNCT TNHPINPKGL ELDPEGSLHH QQKREAPSRS  900
SASSGPQILK CPEAECFRLR CELGPLHQQE SQSLQLHFRV WAKTFLQREH QPFSLQCEAV  960
YKALKMPYRI LPRQLPQKER QVATAVQWTK AEGSYGVPLW IIILAILFGL LLLGLLIYIL 1020
YKLGFFKRSL PYGTAMEKAQ LKPPATSDA
```

SEQ ID NO:293 LBH4 DNA SEQUENCE

Nucleic Acid Accession #:   BC001291
Coding sequence: 44-541 (start and stop codons are underlined)

```
1     11    21    31    41    51
|     |     |     |     |     |
GGGGGCGCCG CGCGCTGACC CTCCCTGGGC ACCGCTGGGG ACGATGGCGC TGCTCGCCTT   60
GCTGCTGGTC GTGGCCCTAC CGCGGGTGTG GACAGACGCC AACGTGACTG CGAGACAACG  120
AGATCCAGAG GACTCCCAGC GAACGGACGA GGGTGACAAT AGAGTGTGGT GTCATGTTTG  180
TGAGAGAGAA AACACTTTCG AGTGCCAGAA CCCAAGGAGG TGCAAATGGA CAGAGCCATA  240
CTGCGTTATA GCGGCCGTGA AAATATTTCC ACGTTTTTTC ATGGTTGCGA AGCAGTGCTC  300
CGCTGGTTGT GCAGCGATGG AGAGACCCAA GCCAGAGGAG AAGCGGTTTC TCCTGGAAGA  360
GCCCATGCCC TTCTTTTACC TCAAGTGTTG TAAAAATTCGC TACTGCAATT TAGAGGGGCC  420
ACCTATCAAC TCATCAGTGT TCAAAGAATA TGCTGGGAGC ATGGGTGAGA GCTGTGGTGG  480
GCTGTGGCTG GCCATCCTCC TGCTGCTGGC CTCCATTGCA GCCGGCCTCA GCCTGTCTTG  540
AGCCACGGGA CTGCCACAGA CTGAGCCTTC CGGAGCATGG ACTCGCTCCA GACCGTTGTC  600
ACCTGTTGCA TTAAACTTGT TTTCTGTTGA TTACCTCTTG GTTTGACTTC CCAGGGTCTT  660
GGGATGGGAG AGTGGGGATC AGGTGCAGTT GGCTCTTAAC CCTCAAGGGT TCTTTAACTC  720
ACATTCAGAG GAAGTCCAGA TCTCCTGAGT AGTGATTTTG GTGACAAGTT TTTCTCTTTG  780
AAATCAAACC TTGTAACTCA TTTATTGCTG ATGGCCACTC TTTTCCTTGA CTCCCCTCTG  840
CCTCTGAGGG CTTCAGTATT GATGGGGAGG GAGGCCTAAG TACCACTCAT GGAGAGTATG  900
TGCTGAGATG CTTCCGACCT TTCAGGTGAC GCAGGAACAC TGGGGGAGTC TGAATGATTG  960
GGGTGAAGAC ATCCCTGGAG TGAAGGACTC CTCAGCATGG GGGGCACGTGG GGCACACGTT 1020
AGGGCTGCCC CCATTCCAGT GGTGGAGGCG CTGTGGATGG CTGCTTTTCC TCAACCTTTC 1080
CTACCAGATT CCAGGAGGCA GAAGATAACT AATTGTGTTG AAGAAACTTA GACTTCACCC 1140
ACCAGCTGGC ACAGGTGCAC AGATTCATAA ATTCCCACAC GTGTGTGTTC AACATCTGAA 1200
ACTTAGGCCA AGTAGAGAGC ATCAGGGTAA ATGGCGTTCA TTTCTCTGTT AAGATGCAGC 1260
CATCCATGGG GAGCTGAGAA ATCAGACTCA AAGTTCCACC AAAAACAAAT ACAAGGGGAC 1320
TTCAAAAGTT CACGAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAA
```

426

SEQ ID NO:294 LBH4 Protein sequence:
Protein Accession #:   AAH01291

```
    1    11    21    31    41    51
    |     |     |     |     |     |
MALLALLLVV ALPRVWTDAN LTARQRDPED SQRTDEGDNR VWCHVCEREN TFECQNPRRC   60
KWTEPYCVIA AVKIFPRFFM VAKQCSAGCA AMERPKPEEK RFLLEEPMPF FYLKCCKIRY  120
CNLEGPPINS SVFKEYAGSM GESCGGLWLA ILLLLASIAA GLSLS
```

It is understood that the examples described above in no way serve to limit the true scope of this invention, but rather are presented for illustrative purposes. All publications, sequences of accession numbers, and patent applications cited in this specification are herein incorporated by reference as if each individual publication or patent application were specifically and individually indicated to be incorporated by reference.